Vol. 58, No. 3 May, 1951

Psychological Review

EDITED BY
CARROLL C. PRATT
PRINCETON UNIVERSITY

CONTENTS

The Current Status of Motivational Psychology ..... SIGMUND KOCH 147

The Brain Analogy ..... H. EDGAR COBURN 155

Two Kinds of Experiment Distinguished in Terms of Statistical Operations ..... MELVIN R. MARKS 179

Response-Selection in Discriminative Learning ..... PHILIP WEISE AND M. E. BITTERMAN 185

Conditioning and Conditionality (Discrimination) ..... O. H. MOWRER AND R. R. LAMOREAUX 196

Idiodynamics in Personality Theory with Special Reference to Projective Methods ..... SAUL ROSENZWEIG 213

Some Theoretical Considerations of Learning ..... W. J. BROGDEN 224

Some Methodological Comments Concerning Expectancy Theory ..... PAUL E. MEEHL AND KENNETH MACCORQUODALE 230

PUBLISHED BI-MONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.
$5.50 volume $1.00 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

The *Psychological Review* is devoted primarily to articles in the field of general and theoretical psychology. This area is obviously difficult to define and delimit, but in view of the large number of manuscripts sent to the editor on all kinds of topics an attempt has to be made to draw the line somewhere.

Ordinarily manuscripts that run to more than about 7500 words are not accepted. This policy is followed partly in an effort to reduce lag of publication and partly from the conviction that brevity which is not inconsistent with clarity is the best way to present an argument.

If an author is prepared to pay for the cost of printing his article, he may arrange for earlier publication without thereby postponing the appearance of manuscripts by other contributors.

Tables, footnotes and references as well as text of manuscripts should be typed double-spaced throughout.

VOL. 58, NO. 3 MAY, 1951

THE PSYCHOLOGICAL REVIEW

THE CURRENT STATUS OF MOTIVATIONAL PSYCHOLOGY [1]

BY SIGMUND KOCH

Duke University

Disorientation is the natural condition of psychology, but the quality and amount of this commodity change at different periods. During the two decades or so before the war, when the tidy trinity of S–R, Gestalt and Psychoanalytic types of theory dominated the scene, disorientation was a property of the whole body psychological, rather than its component parts. Members of each theoretical school had a clear, if over-optimistic, sense of direction; disorientation was a molar product of the "cross-purposeful" relations of each theoretical camp to the others.

With the conclusion of total war, psychology seems now to have entered an era of total disorientation. Pressures deriving from the needs of a sick society, and, more directly, the unavailability, to a greatly expanded corps of psychological and social technologists, of adequate scientific tools for meeting these needs, have led to a widespread reëvaluation of theoretical and fundamental psychology. A sense that the chief pre-war theoretical programs were over-ambitious in scope, prematurely timed in relation to extant empirical knowledge, is beginning to emerge.

I feel that nothing could be, *in principle*, more salutary than this recognition of inadequacy, but the shock of recognition has been a difficult trauma to contain. The messianic impulse rules the day—psychology is to be saved, and the H bomb foiled by group dynamics, cybernetics, transactional observation, projective measurement, unified social science theory, the new look in perception; by demolition of the barricades separating learning, perception and motivation; by the exhaustive analysis of bar-pressing behavior, etc. Even the pre-war messiahs (who shall remain nameless) have lost their old buoyancy. One of them is calling ruefully for "neurological models" despite a long and checkered history of *amours* with "schematic sow bugs" and the like, and others appear to be feeling only half safe, after many years of militant confidence.

One cannot avoid strong concern over the current messianic wave. Most of the very real deficiencies of the synoptic pre-war theories seem generously compounded, not reduced, in the new messages. Apparently the important

[1] The substance of this paper was presented as part of a Symposium on Motivation, held April 6, 1950, at the meeting of the Southern Society for Philosophy and Psychology in Nashville, Tenn. The paper is published exactly as given, except for an altered final paragraph and the inclusion of references. The references are included for illustrative and explicative purposes. They are not intended to provide a comprehensive, or even a representative, bibliography of the literature on motivation.

lesson has not been learned. The lesson is very simple and very painful. It reads: Psychology is not ready for high-order theory of any great range or predictive power. We lack basic areas of empirical knowledge of the sort necessary for adequate theory. We are even a long way from the resolution of many methodological problems, answers to which are crucial to adequate construction of theory. Contrary to certain of the stereotypes of the thirties and forties, we are *not* in a "Galilean era" of theoretical construction; we are closer to Thales than we are to Galileo. We are a science still groping for the identification of our basic variables.

Only *after* learning this lesson of humility will it become possible to make efficient progress towards the objectives of sound theory, and powerful technology. I believe that a sober, rational program for the rôle of theoretical psychology in working towards these objectives can already be blocked out, but this report is not the place for the discussion of such a program.[2] I shall try instead to derive a few of the implications of our "lesson of humility" for the current status of "motivational psychology."

Nowhere is disorientation—old style or new—more evident than in the field of motivation. Every major theory of behavior of the modern period makes some provision for the incorporation of variables which represent motivational processes. Concepts of the most diverse sorts carry this systematic burden in the different theories. They vary in every possible dimension. They vary in level of analysis, from assorted degrees of the molar in behavioral theories, to assorted degrees of the molecular in physiological theories. They vary in *generality* from Freud's *libido* and McDougall's *horme*, to the tension systems which Lewin correlates with his quasi-needs and intentions. To these concepts may be imputed the most widely divergent properties: In some theories they may represent a class of persisting stimuli (*e.g.*, Guthrie, earlier Hull); in others an energy source (Freud, McDougall, instinct theory); in still others a class of regulating and directing mechanisms (Morgan's C.M.S., Köhler's brain-field vectors, Hull's r_G, etc.). Again, there is wide variation in the way in which motivational variables are interrelated with other classes of variables (learning, perceptual, etc.) in the different theories.

To this cacophony of theoretical "bebop," the "new disorientation" has added some new dissonances. Krech, for instance, would save the day by proposing, as the central construct of psychology, a notion of "dynamic systems" which "vary along the dimensions of Isolation, Differentiation and Tension" (19, p. 78). The arbitrarily discriminated fields of motivation, learning and perception immediately wither away, everything now becoming a global matter of dynamic systems. Of course, "variations of . . . Tension will contribute to the conscious correlate of . . . desire, need, demand," and "these particular conscious correlates are the aspects of experience to which most of us refer when we speak about 'motivation'" (19, pp. 78–79). Hochberg and Gleitman, too, have a way out. They "would subsume motivational phenomena under perceptual laws" (12, p. 183).

Confronted with this over-profusion of conflicting "theory," the hygienic

[2] I believe that progress towards adequate theory can only be achieved by substituting for the continued, undisciplined spinning of *comprehensive psychological "theories,"* the pursuit of theoretical psychology, as defined in terms of a set of modest objectives geared to a realistic estimate of the status of our knowledge. I have attempted to outline such a set of *feasible* objectives for theoretical psychology in an article (18) which will appear in this JOURNAL.

thing to do is to turn towards the facts. What we find is a ridiculously meager set of scattered experimental findings and empirical observations. Moreover much of this material proves, on close analysis, to be ambiguous, unreliable, of indeterminate generality, or downright trivial.

The rate of progress in the accumulation of basic empirical relationships has been fantastically slow. Take as a single example of this creeping development the history of the research on the rôle of differential intensities of primary needs or drives in what was called "learning." This, after all, is not an unimportant problem; every theory of behavior must make some commitment with regard to learning-motivation interrelations. In 1918 Szymansky (43) reported that mice learn a simple maze with increasing rapidity as hunger is increased from ½ to 21–24 hours. Learning rate is therefore a positive function of strength of motivation. During the next 15 years or so, a few more experiments of similar design were made, with similar conclusions, despite the fact that Ligon's (24) 6- and 12-hour groups showed an inversion. By the mid-thirties, psychology had finally, via the latent learning experiments, won its way to the distinction between learning and performance. With this distinction nothing was done for another dozen years or thereabouts, except for the prosecution of a violent quarrel over whether latent learning demonstrates the latent death of reinforcement theory. It was, in any case, a well-known reinforcement theorist (14), who, by the early forties, helped psychologists see that the initial question must be broken down into two major problem clusters—one concerning the rôle of drive strength in the *acquisition* of habit structures, and the other having to do with the rôle of drive strength in the utilization (performance) of habit structures already learned. By 1950 we have accumulated some slight evidence, far from quantitatively exact, about the form of the relationship between drive intensity and *vigor* of performance, at least in the case of a single drive (hunger), a single organism (the rat), and a single response (bar pressing) (37, 40). With regard to the effect of variations in drive strength on the rate of trace or habit acquisition, we have not been quite so fortunate. The past decade has yielded some half dozen rat experiments in two standard situations (*e.g.*, 9, 17, 26, 41). Most of these studies show *no effect* of varied food-deprivation periods on the indirect measure of acquisition rate; some, however, *do*.

Now, I submit that it is absurd to proceed with the construction of theories, resembling in any degree the ambitiousness of extant formulations, on the basis of experimental material of this order. The alternative is the drab one of working modestly *towards* theory, and this must be done on a number of fronts. In the area of motivation perhaps the most important need is for a sharp acceleration of progress in the accumulation of "theoretically neutral" empirical knowledge.[3]

Thus I cannot discharge the obligation imposed on me by this symposium in any direct way. The systematic

[3] The expression "theoretically neutral" may require explanation. Without getting twisted in the epistemology of bias, reference frames, etc., I may say this: There is good evidence to suggest that, at those points where the major theoretical programs have been in contact with empirical materials, there has been a growing convergence towards the identification, manipulation and recording of *common experimental variables*, both independent and dependent. Such commonly acknowledged experimental variables may be regarded as "theoretically neutral," and experiments designed to explore relations among such variables may be said to give "theoretically neutral" knowledge.

status of "motivation" cannot be defined by anyone; it can only emerge, along with some future emergence of systematic psychology. What I can do is to examine, briefly, the status of the empirical facts in the field, because the status of motivation is defined by these facts, and not by the empty superstructures of current theory. Since this empirical material, however limited, must provide the starting point for the accelerated development we are calling for, it will be well to present an outline of the way in which the major problem clusters have planlessly evolved.

Any "theoretically neutral" map of the field of motivation must cope with an amorphous and gappy terrain. Any such map must be selective, but my map will try to avoid the kind of selectivity that springs from a desire to chart the field as it should be, rather than as it is. There is no implication that all areas listed are important or indispensable; I imply only that they exist. All that I can squeeze into this report is a map of my map.

As I see it, there are five major problem clusters in the current psychology of motivation. Summarily, these are: I *Functional Problems* (concerning the rôles of needs in behavior), II *Genetic Problems*, III *Problems Concerning Need Interaction and Conflict*, IV *"Mensurational" Problems*, and V *Problems Relating to the Organization of Needs in Personality*.

These areas have emerged independently of coöperative research planning or coördination. They are complexly interrelated, but we do not know how. Advance in any one of these areas often presupposes the application of knowledge that we do not have in some one or more of the others. But this is what we have, and we might as well admit it.

Let us briefly identify these areas.

I. Functional Problems

The problems here are central to the development of behavior theory. The concern is with the general relations between needs and: (a) other behavior determinants (relating to learning, perception, etc.), and (b) final behavior or performance. Functional research has been concentrated in four main subareas:

(1) *The Relations between Needs and Activity* (both general and specific) (32, pp. 368–381, 39). This has long existed as a semi-independent field. In the present state of motivation research, logic would seem to demand the subsumption of this field under the relations between needs and performance.

(2) *The Relations between Needs and Perception* (and allied processes). Here, until recently, we have had only a series of demonstration experiments (2, 4). Currently the sultry glare of the "new look" is producing a flare-up of *conditional* research interest (3, 38, 44) even to the point of a search for "intervening mechanisms." Worthy as such a search may be, my guess is that success in this quest for mediating mechanisms must await *general* advance in psychological theory.

(3) *Relations between Needs and Learning*. Progress in this intricate field has been mainly a matter of achieving the ability to ask *some* of the meaningful questions. We have practically none of the answers. Two principal classes of questions have arisen: The first concerns the rôle of needs in the selective determination of the content of learning (*e.g.*, the extensive literature of "latent learning"). This set of questions is qualitative. The second concerns the rôle of needs in determining the rate of acquisition of learning structures (habits, traces) (9, 17, 26, 41). These questions are, of course, quantitative.

(4) *Relations between Needs and Performance.* This area has also bifurcated into two as yet almost entirely unexplored sets of questions. One of them has to do with the rôle of needs in determining which learning structures are activated under given conditions of stimulation (*e.g.*, 13, 20, and certain of the "latent learning" experiments). The other looks into the effects of differential need intensity on the vigor with which preëxistent learning structures are activated (*e.g.*, 37, 40).

II. Genetic Problems

The genetic research cluster deals with the analysis of: (1) primary need systems, (2) acquired need systems, and (3) the learning mechanisms responsible for the modification of primary and the development of acquired needs.

(1) *Primary Needs.* A mass of research, most of it at the physiological level, and a little at the psychological level, has been accumulated. At the physiological level some progress has been made in the analysis of biological and neural mechanisms associated with primary need states, most conspicuously in the cases of hunger and sex (32, pp. 390–401, 418–438). But what we know about even these needs is microscopic in comparison to what we still have to learn. At the psychological level we have some descriptive material on the behavior tendencies connected with primary needs and their learned patterning. The need for increasing this pathetically limited knowledge has been sharply undervalued.

(2) *Acquired Needs.* Social psychologists, sociologists and cultural anthropologists have provided us with descriptive identifications of major acquired need systems as the latter operate in the statistically defined "individuals" of various cultures and sub-cultures (15, 28). Clinical psychologists and psychiatrists have contributed cross-sectional and longitudinal descriptions of acquired need systems as the latter operate uniquely in individuals (34, 36, 42). I doubt that the specific findings here are of significance to fundamental psychology, although certain of the descriptive and "mensurational" techniques used in these studies *may be.*

(3) *Learning Mechanisms Responsible for Need Development.* The problems in this group are crucial to adequate motivational theory. They are perhaps the most knotty problems that behavior theory must confront. But what do we have? A handful of studies built around a series of vague speculations on (a) "functional autonomy" (1), (b) "response produced stimulation" (6, 30, 31, 33) and (c) "secondary reinforcement" (7, 8, 11, 14).

III. Need Interaction and Conflict

Behavior is never motivated by a single need; it is always a complex function of a plurality of co-acting needs. In this immensely complex area we have virtually no soundly established knowledge. Psychoanalysts and clinicians have given us a few orienting insights and some scattered clinical data. But from experiments we have nothing more than a few surface-scratching studies of irrelevant drive (45), drive combination (16, 17) and conflict (6, 29). Some of these studies come from fundamental psychologists, who, as a group, have shown insufficient interest in these problems; others come from the psychopathologists (25, 27).

IV. Need Measurement

"Measurement" is a metaphor in this connection, as it usually is in psychology. Of course, we have our well-known 86 techniques for estimating the strength and qualitative character of needs in humans and other animals (5,

14, 23, 35), but we have not even begun to tackle those scaling problems, solutions to which are most vital to *fundamental* advance. For instance, we still have not explored the possibility of expressing strength of primary motivation in units which permit translation from one drive to others.

V. Organization of Needs in Personality

This area can only arbitrarily be separated from the general problems of need measurement. Ability to measure or estimate isolated need systems is only a part of what we require, if psychology is to be able to represent its "initial conditions" in a way necessary for prediction. We need also descriptive tools for representing the *interrelations* of needs within the individual. Freud (10) and Lewin (21, 22, 23) (and certain of the personality theorists) have given us some crude, if illuminating, leads in this direction. But they are crude, and they are only leads. The problems of this and the preceding cluster are partly methodological in nature, but concrete solutions will involve much experimentation.

It should be clear by now that what we are mapping can hardly be called "terrain." The subject-matter of our map is still in the gaseous stage, to borrow a geophysical metaphor, and will require much cooling before it assumes determinate form. That cooling can only be precipitated by tough-minded research, and it is redundant to add that the rate of cooling will depend on the rate of research.

There may be a danger in concentrating, as I necessarily have, on only one of the preconditions to sound progress in a field like motivation—the need for an accelerated rate of fundamental research. I hope that my position will not be interpreted as merely another of the indiscriminate invitations to research of the sort connected with the usual laments about the backward condition of the social sciences. If I have asked for research, it is research of a special kind. It is research designed to provide the inductive materials for theory in some non-illusory sense of the word. Thus, it is research guided by a strategy based on an analysis of the prerequisites of adequate theory. The programming of such research must start from an assessment of what is useful and what expendable, what is common and what specific in extant theoretical sallies. It must presuppose a search for those "theoretically neutral" empirical variables which have emerged as common components of the concrete research inspired by our pre-theoretical and quasi-theoretical thinking; it must also presuppose an effort to chart the contours of existing knowledge in terms of the interrelations among such variables. The research I have in mind would be calculated to increase the accuracy, fill in the finer details, and enlarge the boundaries of such "theoretically neutral" knowledge. Yet it would not preclude *conceptual* experimentation with independent theoretical hypotheses or limited-scope postulate sets which have a domain corresponding to some part of this knowledge. Indeed, our fund of theoretically neutral knowledge could not grow without the continuous and intensive prosecution of such theoretical endeavors. One of the main functions of our inventory of theoretically neutral knowledge would be to make attempts at concrete theory-construction feasible, and one of the main results of these concrete, if limited, theoretical attempts would be the enrichment of such knowledge. Only in this way, by continual interplay between restricted, even fragmentary, theoretical attempts, and our inventory of theoretically neutral knowledge, can we purchase our slow way towards theory of a scope and predictive fertility sufficient to justify the business of psychology.

REFERENCES

1. ALLPORT, G. *Personality.* New York: Henry Holt and Co., 1937.
2. ANSBACHER, H. Perception of number as affected by the monetary value of the objects. *Arch. Psychol.*, No. 215, 1937.
3. BRUNER, J. S., & POSTMAN, L. Perception, cognition, and behavior. *J. Personality*, 1949, **18**, 14–31.
4. CARMICHAEL, L., HOGAN, H. P., & WALTER, A. A. An experimental study of the effect of language on the reproduction of visually perceived form. *J. exp. Psychol.*, 1932, **15**, 73–86.
5. CATTELL, R. B. *Description and measurement of personality.* Yonkers: World Book Co., 1946.
6. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy: An analysis in terms of learning, thinking and culture.* New York: McGraw-Hill, 1950.
7. ESTES, W. K. Generalization of secondary reinforcement from the primary drive. *J. comp. physiol. Psychol.*, 1949, **42**, 286–295.
8. ——. A study of motivating conditions necessary for secondary reinforcement. *J. exp. Psychol.*, 1949, **39**, 306–310.
9. FINAN, J. L. Quantitative studies in motivation. I. Strength of conditioning in rats under varying degrees of hunger. *J. comp. Psychol.*, 1940, **29**, 119–134.
10. FREUD, S. *The ego and the id.* London: Hogarth, 1927.
11. GRICE, G. R. The relation of secondary reinforcement to delayed reward in visual discrimination learning. *J. exp. Psychol.*, 1948, **38**, 1–16.
12. HOCHBERG, J. E., & GLEITMAN, H. Towards a reformulation of the perception-motivation dichotomy. *J. Personality*, 1949, **18**, 180–191.
13. HULL, C. L. Differential habituation to internal stimuli in the albino rat. *J. comp. Psychol.*, 1933, **16**, 255–273.
14. ——. *Principles of behavior.* New York: Appleton-Century, 1943.
15. KARDINER, A. *The psychological frontiers of society.* New York: Columbia Univ. Press, 1945.
16. KENDLER, H. H. Drive interaction. I. Learning as a function of the simultaneous presence of the hunger and thirst drive. *J. exp. Psychol.*, 1945, **35**, 96–109.
17. ——. Drive interaction. II. Experimental analysis of the role of drive in learning theory. *J. exp. Psychol.*, 1945, **35**, 188–198.
18. KOCH, S. Theoretical psychology, 1950: An overview. PSYCHOL. REV., 1951, **58**. (In press.)
19. KRECH, D. Notes toward a psychological theory. *J. Personality*, 1949, **18**, 66–87.
20. LEEPER, R. The role of motivation in learning: A study of the phenomenon of differential motivational control of the utilization of habits. *J. genet. Psychol.*, 1935, **46**, 3–40.
21. LEWIN, K. *Dynamic theory of personality.* New York: McGraw-Hill, 1935.
22. ——. *Principles of topological psychology.* New York: McGraw-Hill, 1936.
23. ——. The conceptual representation and the measurement of psychological forces. *Contr. Psychol. Theory*, Vol. 1, No. 4. Durham, N. C.: Duke Univ. Press, 1938.
24. LIGON, E. M. A comparative study of certain incentives in the learning of the white rat. *Comp. Psychol. Monogr.*, 1929, **6**, No. 2, 1–95.
25. LURIA, A. R. *The nature of human conflicts.* New York: Liveright, 1932.
26. MACDUFF, M. N. The effect on retention of varying degrees of motivation during learning in rats. *J. comp. Psychol.*, 1946, **39**, 207–240.
27. MASSERMAN, J. H. *Behavior and neurosis.* Chicago: Chicago Univ. Press, 1943.
28. MEAD, M. *Cooperation and competition among primitive people.* New York: McGraw-Hill, 1937.
29. MILLER, N. E. Experimental studies of conflict. In *Personality and the behavior disorders* (J. McV. Hunt, Ed.). New York: Ronald Press, 1944. Vol. 1.
30. ——. Studies of fear as an acquirable drive. I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, **38**, 89–101.
31. ——, & DOLLARD, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
32. MORGAN, C. T. *Physiological psychology* (2nd ed.). New York: McGraw-Hill, 1950, Chap. 17.
33. MOWRER, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. PSYCHOL. REV., 1939, **46**, 553–565.
34. MULLAHY, P. (Ed.). *A study of interpersonal relations.* New York: Hermitage Press, 1949.
35. MURPHY, G. *Personality.* New York: Harper & Bros., 1947.
36. MURRAY, H. A. *Explorations in personality.* New York: Oxford Univ. Press, 1938.

37. PERIN, C. T. Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93–113.
38. PROSHANSKY, H., & MURPHY, G. The effects of reward and punishment on perception. *J. Psychol.*, 1942, **13**, 295–305.
39. REED, J. D. Spontaneous activity of animals: A review of the literature since 1929. *Psychol. Bull.*, 1947, **44**, 393–412.
40. SALTZMAN, I., & KOCH, S. The effect of low intensities of hunger on the behavior mediated by a habit of maximum strength. *J. exp. Psychol.*, 1948, **38**, 347–370.
41. STRASSBURGER, R. C. Resistance to extinction of a conditioned operant as related to drive level at reinforcement. *J. exp. Psychol.*, 1950, **40**, 473–487.
42. SULLIVAN, H. S. *Conceptions of modern psychiatry*. Washington, D. C.: The William Alanson White Psychiatric Foundation, 1947.
43. SZYMANSKI, J. S. Ein experimenteller Beitrag zur Analyse der bei Entstehung neuer Gewohnheiten mitwirkenden Faktoren. *Pflüg. arch. ges. Physiol.*, 1918, **170**, 197–220.
44. WALLACH, H. Some considerations concerning the relation between perception and cognition. *J. Personality*, 1949, **18**, 6–13.
45. WEBB, W. The motivational aspect of an irrelevant drive in the behavior of the white rat. *J. exp. Psychol.*, 1949, **39**, 1–14.

[MS. received July 6, 1950]

THE BRAIN ANALOGY [1]

BY H. EDGAR COBURN

Registered Civil Engineer, San Diego, California

INTRODUCTION

The Brain Analogy is a theory of behavior, an instrument for psychological and physiological research, and a specification for the design and construction of mechanisms possessing intelligence.

The basis of this theory is a hypothetical nervous system inferred primarily from data on behavior experiments and to a lesser extent from anatomical data. It proposes to simulate function, not structure, of organic nervous systems. This hypothetical nervous system, or Brain Analogy (BA), has structures and properties assigned to it as indicated in the postulates to follow. The sole criterion in the selection of these postulates has been to achieve a *useful* result, a *workable* instrument for research.

The following tentative definition of *intelligence*, based upon this research, is used throughout the paper.

Intelligence is the property of reacting on a basis of probability as determined by the individual organism's incomplete sensory samples of the environment.

The fundamental premise of science is that the universe is ordered and that research can reveal that order. A sample of part of the universe of course leaves the door open for erroneous conclusions to enter and the best guess may be the most probable one. If the organism sampled the entire universe there would be no necessity for intelligence. Whenever the sampling technique gives a probability of less than unity, or when the "facts" are too numerous to cope with, the organism needs a nervous system that functions on a basis of probability. When the probability is unity and the facts are not too numerous, intelligence is not only useless, it is positively a handicap because certainty and a simpler mechanism achieve the desired result in an unequivocal manner. The mechanism of intelligence is always subject to deflections by random stimuli and fluctuating thresholds.

The mechanism of conditioning must be arranged to favor the most probable hypothesis. The strongest component of a compound stimulus should get most of the conditioning (5, p. 142). Learning needs to be general at first (5, p. 113) and specific later. Signals near in time (but not too near) are to be favored and signals separated further from reinforcement (delayed and trace reflexes) should receive less conditioning, while signals which occur after reinforcement are to receive only weak and transitory conditioning (5, p. 393), if any. These conditions are fundamental and are unchanged by the fact that selective reinforcement will raise a weak component to a position of dominance; this acquired dominance is concerned merely with the magnitude of the conditioned response after training and not with the relative rate of change of conditioning.

Although the BA is quantitatively unlike any species, it serves to elucidate principles which apply to nervous systems in general, and should aid in research by limiting the number of possibilities that need be considered and investigated. While the BA is, or can be, a definite physical structure pos-

[1] The author wishes to express his gratitude to Vera Jane Coburn for her assistance in preparing the manuscript and drawings.

sessing properties that can be investigated critically, both by theory and measurement, and compared exhaustively with data from experiments with living organisms, the tremendous task of accurate detailed alignment is not the goal of this project—the primary object of this paper is the establishment of the *principles* of intelligent mechanisms.

POSTULATES

The following postulates are general and contain little of a quantitative nature. Where it seems helpful the postulates are accompanied by a discussion of the factors that led to their selection. Where it seems necessary, resort is had occasionally to the calculus, but the graphs and common algebra suffice to give an adequate grasp of the theory.

Figure 1 shows some of the relationships defined in the postulates and indicates the general structure of the BA.

Postulate 1

A sensory neuron (in BA theory) is a unit consisting of one or more receptor elements or poles, latency elements, a combining element, and a conducting element which divides into a multitude of branches each terminated with a conditioning element (delta cell).

Postulate 2

A motor neuron is a unit consisting of one or more receptor elements or poles, a combining element, a conducting element, a cell body to which sensory neurons make potential connections, a drive receptor having a separate connection with the cell body, and an output element which actuates an effector.

Postulate 3

A neuron responds on an all-or-none basis; conduction is unidirectional.

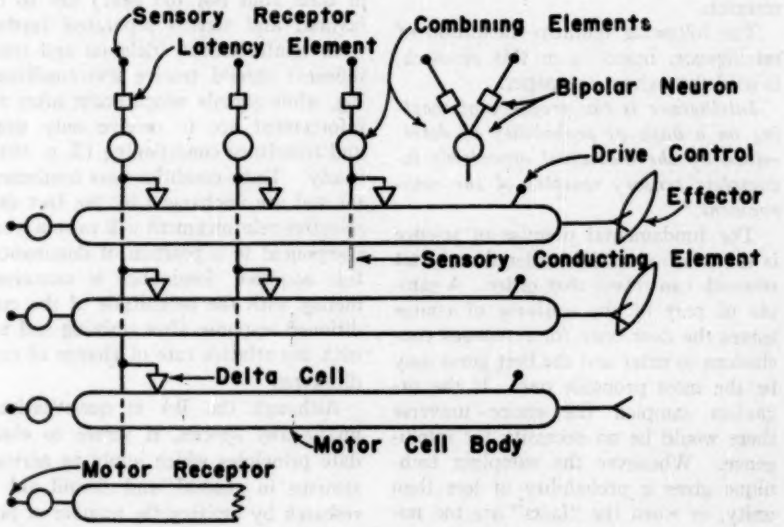

FIG. 1. The structure of the Brain Analogy is shown as it pertains to the supposed functions of the cerebral cortex. Subcortical functions have been omitted. No significance attaches to the serial order of the delta cells on the motor cell bodies though there is a proximity effect.

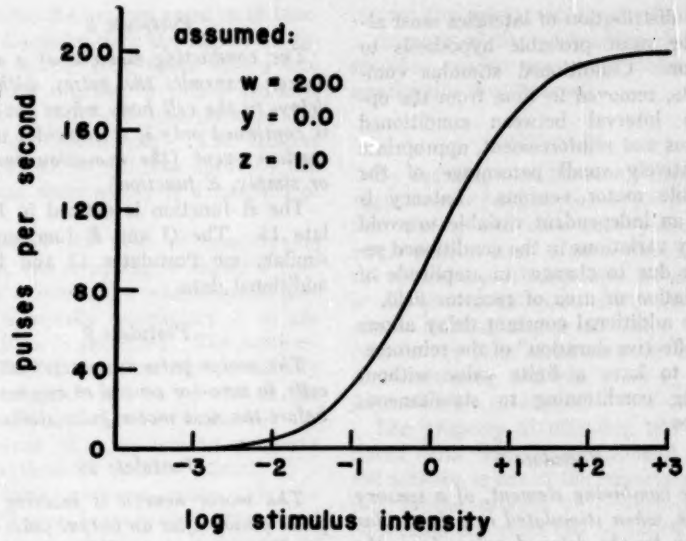

FIG. 2. Graph of Equation 1.

Postulate 4

A receptor element or pole, when appropriately stimulated, responds with a pulse frequency that is a function of stimulus strength. The response rate in pulses per second is

$$P = \frac{w}{z\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(x-y)^2}{2z^2}} dx, \quad (1)$$

where w, y, and z in a population of receptors are independent variables with normal distribution. For any one receptor they are constants.

P = *pulses per second,*
e = *the base of natural logarithms,*
w = *saturation value of pulse rate,*
x = *log intensity of stimulus,*
y = *log mean-pulse-frequency stimulus,*
z = *standard deviation of pulse distribution.*

There is no finite threshold for receptors: any stimulus greater than zero results in an output frequency. The function y is the substitute for the conventional threshold value but since z and w vary from receptor to receptor it cannot offer a direct comparison. A finite threshold has some advantages, if expressed as a satisfactory mathematical function, but may not be essential from an operational viewpoint. Figure 2 is the graph of Equation 1.

Postulate 5

The pulses generated by the receptor elements of the sensory neurons are delayed during conduction to the combining element; the latency distribution is

$$L = \frac{1}{\sigma_1\sqrt{2\pi}} e^{\frac{-(T-T_0)^2}{2\sigma_1^2}} \quad (2)$$

where

k = *a constant delay for all receptors,*
L = *probability density of latencies,*
t = *time in seconds (not less than k),*
T = *log $(t - k)$,*
T_0 = *mean value of T,*
σ_1 = *standard deviation of the latency distribution.*

The distribution of latencies must allow the most probable hypothesis to function. Conditioned stimulus components, removed in time from the optimum interval between conditioned stimulus and reinforcement, appropriate a relatively small percentage of the available motor neurons. Latency is made an independent variable to avoid latency variations in the conditioned response due to changes in amplitude of stimulation or area of receptor field.

The additional constant delay allows the "effective duration" of the reinforcement to have a finite value without causing conditioning to simultaneous reflexes.

Postulate 6

The combining element, of a sensory neuron, when stimulated in a particular fashion by the delayed pulses from the poles, generates a pulse frequency that is a function of the stimulus strength.

There is nothing in this postulate that limits the relationship of one pole to another or to the combining element. The multipolar neurons may require stimulation of all the receptors in order to respond. The relative magnitude of the stimulus for the several poles may be set at any finite value, or substantially zero if required, and the response is not forthcoming unless the stimuli strengths or strength ratios correspond. The poles may belong to separate sensory fields or analysers, and the latencies may take any useful values. Special combinations of properties, including generalized sensory neurons, will be discussed later.

Postulate 7

Each sensory neuron has one potential connection with every motor neuron cell body by means of the delta cells which receive the sensory pulses transmitted through the branches of the conducting element.

Postulate 8

The conducting element of a motor neuron transmits the pulses, with zero delay, to the cell body where the pulse is continued only if it coincides with a random event (the evocation function or simply, E function).

The *E* function is defined in Postulate 15. The *Q* and *E* functions are similar; see Postulates 12 and 14 for additional data.

Postulate 9

The motor pulse "contacts" all delta cells, in turn (or several at any instant), before the next motor pulse starts.

Postulate 10

The motor neuron is inactive for a finite period after an output pulse either conditioned or unconditioned.

See Postulate 14 for related data.

Postulate 11

The coincidence of a sensory pulse, motor pulse, and a random event (the Q function) at the delta cell (reinforcement) changes the potential connection into a functional one (conditioning) unless the cell is disabled.

Postulate 12

There exists a factor (the Q function) such that the rate of conditioning for a group of neurons is independent of the strength of sensory stimuli, is reduced by massed trials, is increased by drive, and is substantially independent of the relative location of the delta cells on the motor cells.

The reader may, if he wishes, consider that the motor neuron cell body contains a cloud of molecules moving about in random fashion. The cloud consists of two components: molecules inert toward the conditioning reaction and molecules, called *ignitors*, the presence of one of which at the delta cell

constitutes the random event or Q function. Consider that the ignitor is destroyed by the coincidence with a sensory pulse whether a motor pulse is present or not and regardless of the status of conditioning. The ignitors are replaced by diffusion into the cell body and their supply is a function of drive.

Whatever the facts of conditioning may be, the foregoing is merely a convenient fiction and a mathematical postulate is equally satisfactory if an abstract form is preferred. The randomness of the Q function (the probability that an ignitor will be present at the delta cell) is required to prevent the appearance of a systematic property with rhythmic or periodic stimuli.

A simple analogue exists (Fig. 3) that will serve in place of equations and which seems useful as a component in electronic models of the BA. The condenser is charged to the voltage of the battery through the resistor R_1. The resistor R_2 is continuously maintained inversely proportional to the total pulse rate applied to the motor neuron through the delta cells. When the pulse rate declines again to zero the condenser regains its charge slowly so as to provide for a massed-trials effect. At any instant the voltage on the condenser is a measure of the probability that the specified random event will occur in a small finite time interval.

Postulate 13

By inducing into the motor neuron a negative phase which persists for a finite time during and after each pulse, the motor neuron itself after a finite interval of activity (the effective duration), or the first conditioned delta cell into action disables all, or all other, delta cells on the same motor neuron, insofar as conditioning and the effect of reinforcement are concerned.

The property of disabling the more latent delta cells, by conditioned delta cell activity, is one of the important factors in the process of differentiation. It permits one delta cell to appropriate the motor neuron and hold it against the encroachments of all other cells under a given set of conditions. After one delta cell has become conditioned, on subsequent trials, only cells less latent can compete for conditioned connections. The least latent *of those regularly reinforced* appropriates the motor neuron sooner or later and thereby causes the latency of the conditioned

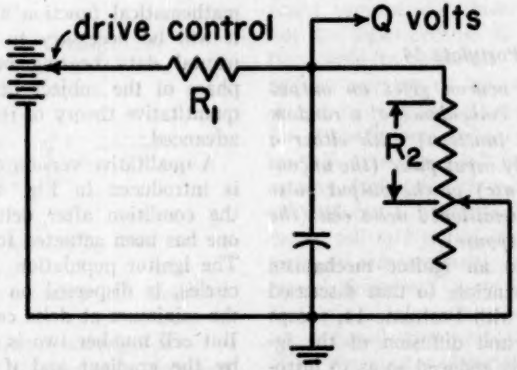

FIG. 3. An electrical analogue for defining magnitude of the Q function probability.

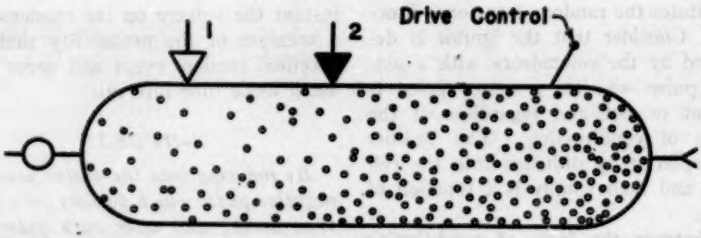

FIG. 4. A gradient is established in the population density of "ignitor" molecules due to the action of delta cell number one. Testing with conditioned cell (2) would now reveal a less-than-normal response which has usually been interpreted as inhibition.

reflex to reach a minimum. Since other cells may also be active and conditioned (redundant conditioning), the disabling property by preventing the effect of reinforcement causes these more latent cells to drop their conditioned status after a sufficient number of conditioned output pulses (see Postulate 17). Conditioning is established on a competitive basis, the successful competitor depending on a combination of receptor properties and the experimental program. Actually, the properties of the delta cells are involved too, if differentiation is based on the decrement factor.

The motor-induced negative phase prevents backward conditioning.

There is nothing in this postulate that concerns conditioned but inactive delta cells.

Postulate 14

The motor neuron gives an output pulse for each coincidence of a random event (the E function) with either a motor cell body input pulse (the unconditioned response) or the output pulse of an active conditioned delta cell (the conditioned response).

This is also an ignitor mechanism identical in principle to that discussed in connection with Postulate 11, except that motility and diffusion of the ignitors is greatly reduced so as to introduce a pronounced temporary reduction of ignitor population in the area surrounding the center of activity. Other conditioned delta cells activated in the area of reduction show less than normal response (the proximity effect).

Postulate 15

There exists a factor (the E function) such that the pulse rate of conditioned responding is reduced by proximity on the motor cell to another group of active, or recently active, delta cells, by massed test trials, and by reduction in drive.

This property is responsible for such characteristics as external inhibition, irradiation, uninterrupted extinction, and spontaneous recovery.

It may not be possible to express Postulate 15 in the form of an elementary mathematical function and in addition it will be necessary to collect experimental data bearing on this specific phase of the subject before any valid quantitative theory of responses can be advanced.

A qualitative version of Postulate 15 is introduced in Fig. 4 which shows the condition after delta cell number one has been actuated for a short time. The ignitor population, represented by circles, is dispersed on gradients with the minimum at delta cell number one. But cell number two is also influenced by the gradient and if it is a conditioned cell and actuated at the stage

shown in the diagram the conditioned response is less than normal. The rates of decline and recovery are influenced by the motility of the ignitors and by the rate of diffusion through the cell wall.

Postulate 16

Drive, the effect of physiological conditions on the nervous system, is one factor controlling the magnitude of the Q and E functions.

The absence of hunger, for example, and attendant disappearance of the Q and E functions makes it impossible to establish a conditioned salivary reflex, or to evoke a response from a previously established one. In an exactly parallel manner the locomotion system is not sensitized during satiation (it may be by some other drive, however), and instrumental conditioning and responding in the conventional sense are impossible based upon hunger.

Postulate 17

After a finite number of conditioned output pulses (the gamma phase), cumulative with separate trials, and subsequent to the last effective reinforcement pulse, each conditioned output pulse exposes the delta cell to a probable loss of conditioning (the decrement factor). The distribution of the decrement factor is

$$D = \frac{M}{\sigma_i \sqrt{2\pi}} e^{\frac{-(v - v_i)^2}{2\sigma_i^2}}, \qquad (3)$$

where

$$v_i = a - \frac{b}{\sigma_i \sqrt{2\pi}} \int_{-\infty}^{v} e^{\frac{-(y - y_0)^2}{2\sigma_i^2}} dy,$$

and

a = *maximum value of* v_i,
b = *range of* v_i,
D = *probability density of decrement factor*,
i = (*see Equation 5 and Fig. 5*),
K = *reciprocal of maximum decrement factor*,
M = *number of delta cells*,
σ_i = *standard deviation of* v,

$$v = \log\left(\frac{1}{\theta} - K\right),$$

v_i = *mean value of* v (*a function of* y),
y_0 = *mean value of* y,

$$\theta = \text{decrement factor} \left(\theta > \frac{1}{K}\right).$$

Gamma phase in terms of the number of conditioned output pulses is

$$\gamma = \frac{J}{\theta}, \qquad (4)$$

where

γ = *gamma phase*,
J = *gamma phase constant*.

It should be noted at once that while the loss of conditioning is based upon probability and is uncertain for any specific cell, the "finite number of pulses," gamma, is a definite number for any cell based upon another definite number—the decrement factor—the *measure* of probability.

Since a single effective reinforcing pulse by re-initiating the gamma phase, at any time prior to extinction, re-establishes the conditioned delta cell to its status of maximum potential resistance to loss of conditioning it might appear that a trivial reinforcement could prevent decrement, but this is not the case because it is practically impossible to stimulate weakly a significant number of motor neurons. It should be observed that reinforcement for a *conditioned* delta cell consists of an "effective duration" motor pulse occurring during the time that the delta cell induces a negative phase into the motor cell and *is not the same mechanism as reinforcement of an unconditioned delta cell* (see Postulate 11). Not only is the mechanism different but one is based upon probability while the other is a matter of certainty (for delta cells possessing appropriate properties).

The introduction of the probability element as a property of decrement causes the organism *gradually* to become immune to occasional non-reinforced trials and a mechanism is provided for the phenomena of periodic reconditioning. Decrement is inversely related to gamma to increase the "reflex reserve" (7, p. 294).

The term *periodic reconditioning* is used in this paper to include any series of trials in which non-reinforced trials intervene with reinforced trials. This, of course, includes as a special case Skinner's use of the term (7), and also his examples of reinforcement at a constant ratio.

The requirement for *conditioned* output pulses introduces a massed-trials effect due to the effect of the E function. Massed non-reinforced trials bring about rapid overt extinction but the permanent extinction is less than with spaced trials.

Glandular reflexes must be arranged so as to avoid useless secretion; the continued action as with musculature reflexes in response to periodic reconditioning, can serve no useful purpose. It is therefore necessary to avoid differentiation based upon low decrement factors and this, though not indicated in equations, is effected by restricting the range of distribution.

The decrement factor appears to be the weakest of all the postulates. It seems probable that a constant or linear function is inadequate and the present choice is admittedly workable over only a limited range, if at all. It is conjectured that the decrement factor must be a function of threshold (or y) in order to provide stability for the magnitude of the conditioned response under the influence of periodic reconditioning; otherwise, delta cells having a high pulse rate will be discriminated against and the conditioned response will decline in strength (E remains unchanged). Evidence to support the conjecture of stability of magnitude of response is scanty and unimpressive when the burden is considered; in fact, there seems to be some biological advantage if the response increases in magnitude.

FUNDAMENTALS

Our next concern is with the application of the basic postulates to simple conditioning, extinction, and differentiation.

An intact living organism may be subjected to an experimental program designed to show habit growth, but inevitably it is also subject to habit extinction and differentiation since it contains the complete intelligence mechanism. The BA presents the same complication. To some extent we must simplify to keep the problem manageable even if this can be done only at the expense of distortion.

An attempt was made to deduce from the postulates the general laws of habit growth and decay. The author was reluctantly forced to the conclusion, however, that a logically complete analysis might require a major effort in itself. Studies were also made of the possibility of revealing the detailed progress of habit growth and decay. The algebra of simple probability seems to apply but it is extraordinarily tedious. While the detailed course of habit growth is thus not delineated and the general laws are not deduced, there is a generality concerning the ultimate fate of a habit that is of such importance for understanding as to deserve the utmost attention. Reference is made to the fact that some conditioned connections possess stability and some do not. This property provides a logical basis for certain simplifications which will shortly allow us to show applications of the BA to the much more complex problems that have engaged the interest of psychologists.

Receptor Fields

Let us assume the existence of two receptor fields, C and U, containing the sensory and motor receptors, respectively, of BA theory. Field C will be the domain of the conventional conditioned stimulus while Field U will be the domain of the conventional unconditioned stimulus or reinforcement. The motor receptors actuate the motor neurons and the unconditioned response. The delta cells of the sensory neurons compete for conditioned connections to the cell bodies of Field U motor neurons.

In this discussion we assume that all neurons have simple one-pole receptor arrangements without qualitative differences within one field.

Field C contains, say, 1,000 sensory neurons while Field U contains 10,000 motor neurons. There will be

$$1{,}000 \times 10{,}000 = 10{,}000{,}000 \text{ delta cells}$$

in accordance with Postulate 7, and there will be 1,000 potential conditioned connections to each motor neuron.

The total pulse rate output of a receptor field is a function of several independent variables: number of receptors, "threshold," deviation of individual response curves, saturation value, and stimulus intensity. Since at this stage we are interested primarily in developing the principles involved in the BA, and the rôle of all of the variables appears not to be known, considerable effort will be saved by assigning constants for some of the variables. It will suffice to retain "threshold" and stimulus intensity as independent variables. The variable y is substituting for the conventional threshold.

The pulse rate of response of a receptor in pulses per second is determined by Postulate 4 where $x = \log I$; and w, y, and z are independent variables to which we assign the arbitrary constants of 200 pulses per second saturation value, a set of 14 discrete values, and unity, respectively, as shown in Fig. 5. It follows that the pulse rate of a field of receptors is

$$\sum P = n \sum_{i=1}^{14} \frac{w\Delta_i}{z\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(x-y_i)^2}{2z^2}} dx, \quad (5)$$

where

Δ_i = fraction of receptors having one of the assigned "threshold" values,
i = an index for the set of constants,
n = number of receptors.

The family of curves shown in Fig. 5 represents a field of receptors with normal distribution of "threshold," or values of mean frequency stimuli. The dashed curve in the same figure is the average response for the group and shows graphically how the useful range of stimulation is increased by a non-homogeneous population of receptors. The ordinates of the dashed curve can be multiplied by n to get the total output pulse rate for the field at any stimulus intensity.

Assume that the 1,000 sensory receptors of Field C are subject to a stimulus intensity of unity ($x = 0$). The corresponding ordinate

$$\frac{\sum P}{n} = 38 \text{ pulses per second}$$

and

$$\sum P = 38{,}000 \text{ pulses per second.}$$

Next consider that the 10,000 receptors of Field U are subject to a stimulus intensity of ten ($x = 1$). As before

$$\frac{\sum P}{n} = 86 \text{ pulses per second}$$

and

$$\sum P = 860{,}000 \text{ pulses per second.}$$

Effect of Latency Elements

Assuming that the outputs of the receptor fields have been computed or

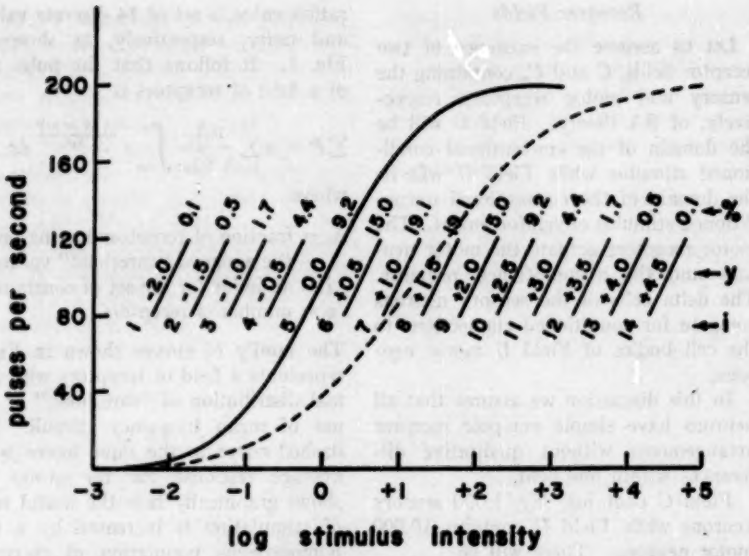

FIG. 5. A family of response curves generated by giving the "threshold" (y) normal distribution. The dotted curve is the average for the group.

measured, the next concern is with the effect of the latency elements on the pulse populations traversing them. The undelayed output of a receptor field is some constant finite value for any given stimulus intensity; but it is necessary to know how the total pulse rate appears at the delta cells if we wish to evaluate the Q function and analyse the mechanism in greater detail. Referring to Postulate 5 we find that the latency is made up of two parts, one of which is constant for all latency elements of sensory neurons. This value, k, is now assumed to be 0.3 second. The latency distribution of Postulate 5 is integrated to determine the fraction of neurons (one receptor per neuron) past latency:

$$\frac{1}{\sigma_1\sqrt{2\pi}}\int_{-\infty}^{T} e^{\frac{-(T-T_0)^2}{2\sigma_1^2}}\,dT. \qquad (6)$$

By assigning for T_0 and σ_1 the values of $\log 0.3 = 9.477 - 10 = -0.523$, and unity, respectively, we get the graph shown in Fig. 6. The ordinate of this curve multiplied by N (the number of neurons stimulated) is the number of neurons past latency for any time t after the beginning of the stimulus.

Since latency is an independent variable the neurons past latency are representative of the entire field insofar as other variables are concerned and therefore the latency curve is a total-pulse-rate curve provided appropriate units are chosen.

For convenience the stimulus applied to the field is uniformly intense at all points and is applied and removed instantaneously with a finite interval between. Mathematically, we apply a positive step function followed after a finite interval by an equal negative step function. The two combined will be termed an *impulse function* (Fig. 7).

It will be apparent at once from the form of the latency equation that some

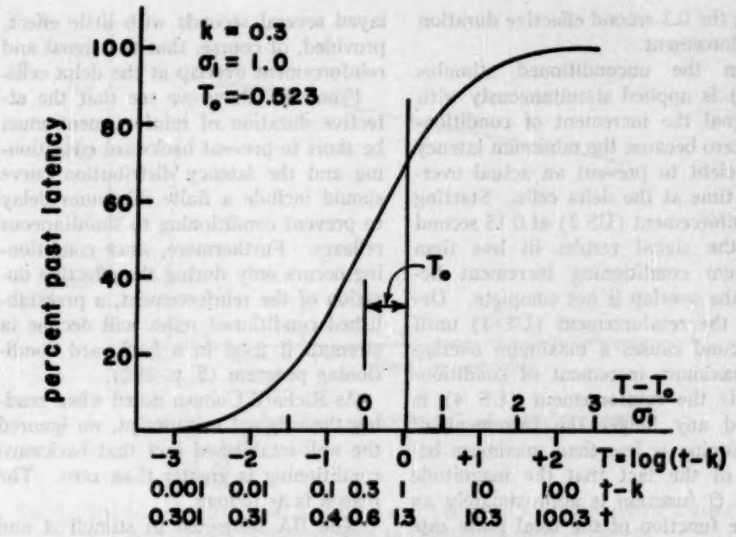

FIG. 6. Percentage of sensory neurons past latency at any time after the beginning of the stimulus.

of the pulses originating with the positive step function will be delayed indefinitely long so that some of the impulse function pulses will have run their course and terminated before others have appeared at the delta cells. A graph of the phenomenon is shown in Fig. 8.

In the typical classical conditioning experiment the conditioned stimulus is applied alone for several seconds followed by the simultaneous application of the conditioned stimulus and reinforcement for several seconds longer. These additional data are also indicated in Fig. 8. The reader is referred to Postulate 13 for the significance of the "effective duration" of the reinforcement which is here assumed to be 0.3 second.

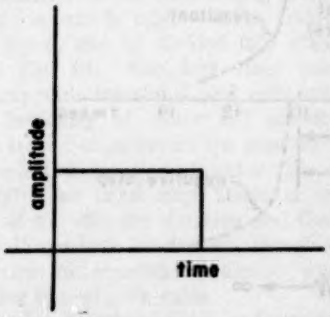

FIG. 7. The impulse function, one of many possible driving functions, is typical of several classical signals.

Stimulus Timing

We propose here to indicate only in a general qualitative way the effect of the time relationship between the conditioned stimulus and reinforcement.

Assume that the conditioned stimulus is a "click" of sufficient intensity and short enough duration to cause substantially all of the receptors to respond simultaneously with one output pulse. The pulse distribution at the delta cells will be a reproduction of the latency curve which is repeated in Fig. 9. The same constants are used as before in-

cluding the 0.3 second effective duration of reinforcement.

When the unconditioned stimulus (US 1) is applied simultaneously with the signal the increment of conditioning is zero because the minimum latency is sufficient to prevent an actual overlap in time at the delta cells. Starting the reinforcement (US 2) at 0.15 second after the signal results in less than maximum conditioning increment because the overlap is not complete. Delaying the reinforcement (US 3) until 0.3 second causes a maximum overlap and maximum increment of conditioning. If the reinforcement (US 4) is delayed any longer the increment of conditioning is less than maximum because of the fact that the magnitude of the *Q* function is approximately an inverse function of the total pulse rate while the pulse rate is increasing, but fails to recover rapidly as the pulse rate declines.

In the event that we use a steady tone as conditioned stimulus instead of a click, the recovery of the *Q* function is an insignificant item and the start of the reinforcement may be delayed several seconds with little effect, provided, of course, that the signal and reinforcement overlap at the delta cells.

From the above we see that the effective duration of reinforcement must be short to prevent backward conditioning and the latency distribution curve should include a finite minimum delay to prevent conditioning to simultaneous reflexes. Furthermore, since conditioning occurs only during the effective duration of the reinforcement, a preëstablished conditioned reflex will decline in strength if used in a backward conditioning program (5, p. 392).

As Richard Coburn noted when reading the original manuscript, we ignored the well-established fact that backward conditioning is greater than zero. The reason is as follows.

The BA subjected to stimuli *A* and *B* is an abstraction having no parallel among living organisms either in or out of the laboratory. Always there are random stimuli to contend with—stimuli not detected by the experimenter—and impossible to eliminate for intact organisms. Those stimuli compete for control of conditioned reflexes and,

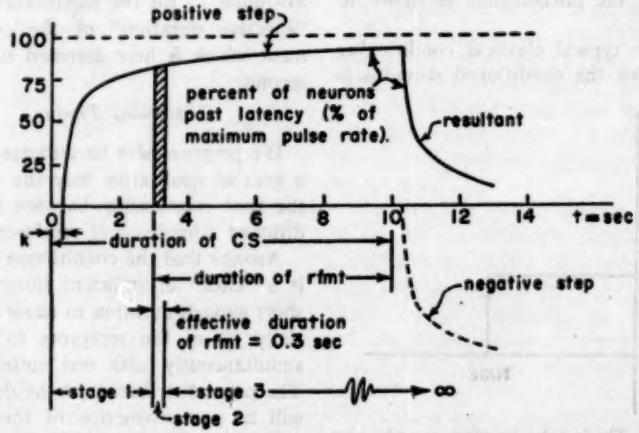

FIG. 8. The percentage of neurons past latency is shown on a linear time scale. The shaded area represents the "effective duration" of the reinforcement.

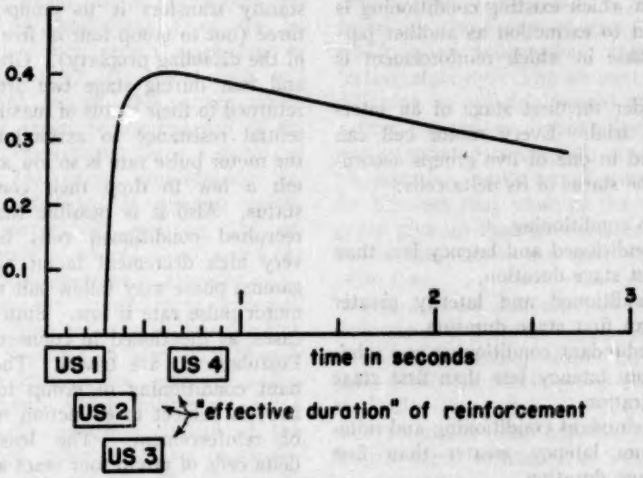

FIG. 9. The latency distribution is plotted to an arithmetic scale with several degrees of asynchronism of reinforcement to indicate the effect of timing.

since some are actuated *before* the conventional conditioned stimulus and reinforcement, no problem of backward conditioning exists for them. Were they not random they would soon appropriate all actuated motor neurons; as it is they acquire control of approximately ten per cent, to judge from experimental data.

Simple Conditioning

Each trial of the conditioning process itself, in which reinforcement follows the signal, can be divided into stages (see Fig. 8). The first stage commences with the signal and ends with the beginning of motor cell activity. The second stage covers the interval of the effective duration of motor cell activity. The third stage starts at the end of the effective duration and theoretically persists indefinitely, though in practice the quantitative aspects soon decline to negligible value.

During the first stage, and we are considering one of the intermediate trials of a series, the delta cells become active as their latent periods expire, exposing the conditioned ones to decrement and establishing negative phases which disable unconditioned and redundant cells. It should be emphasized, however, that in no stage is a cell exposed to decrement until its gamma phase is depleted.

All conditioning takes place during the second stage. Previously-conditioned active delta cells are reinforced, unless they are redundant, in which case they continue to be subject to decrement. See Postulates 11 and 17 concerning conditioning.

In the third stage the negative phase is self-induced in the motor cells and *all* conditioned active delta cells are subject to decrement. The computation of the decrements occurring in this stage is relatively simple except for the fact that the E function is difficult to evaluate.

By confining our attention to an intermediate trial we get a qualitative view of the general case which applies also to the first trial as a particular in-

stance in which existing conditioning is zero, and to extinction as another particular case in which reinforcement is zero.

Consider the first stage of an intermediate trial. Every motor cell can be placed in one of five groups according to the status of its delta cells:

1. No conditioning,
2. Conditioned and latency less than first stage duration,
3. Conditioned and latency greater than first stage duration,
4. Redundant conditioning and minimum latency less than first stage duration,
5. Redundant conditioning and minimum latency greater than first stage duration.

Of course if all trials are alike no delta cell having a latency greater than stage one plus stage two can be conditioned because the effective duration is then past, and such cells are excluded from the discussion. Note that for groups four and five redundancy is not limited to one extra delta cell; and while the least latent conditioned cell in group five is not active in the first stage, it can be active in group four where only one delta cell can have a latency less than the duration of stage one. This is true because the negative phase starts before the reinforcement and prevents further conditioning.

No conditioning takes place during stage one; hence we are concerned only with decrement, and remembering (Postulate 17) that decrement can take place only when a delta cell is giving conditioned output pulses, we eliminate from consideration groups one, three, and five.

Considering now the second stage of an intermediate trial we see that the same five groups previously discussed apply. A group one cell is subject to conditioning which, if accomplished, instantly transfers it to group two or three (not to group four or five because of the disabling property). Groups two and four during stage two are always returned to their status of maximum potential resistance to extinction unless the motor pulse rate is so low as to permit a few to drop their conditioned status. Also it is possible that newly recruited conditioned cells having a very high decrement factor and short gamma phase may follow suit when the motor pulse rate is low. Both of these cases, as mentioned in connection with Postulate 17, are trivial. The redundant conditioning of group four (and five) is subject to extinction regardless of reinforcement. The least latent delta cells of group four react as if they were in group two. The conditioned cells of group three and the least latent conditioned cells of group five are returned to their status of maximum potential resistance to extinction, as for groups two and four, unless conditioning intervenes during the latent period when they are transferred to the redundant class.

Further detailed discussion on a qualitative basis is unprofitable so we will proceed at once to consider the ultimate fate of the conditioning process.

The Physiological Limit

The following simplified approach is based largely upon a non-mathematical argument. Assuming the experimental program shown in Fig. 8, and also assuming the individual trials are spaced far enough in time to avoid interaction due to depression of the E and Q functions, it will be observed that some conditioned connections possess stability and some do not. No delta cell having a latency greater than stage one plus stage two can become conditioned because the effective duration is then past. The balance of the delta cells are subjects for possible con-

ditioning which will be stable only under certain conditions: First, the latency must not exceed stage one or displacement by less latent cells will occur, and second, the gamma phase must not expire or the cell will be exposed to decrement. *Once a delta cell meeting the above requirements becomes conditioned it remains conditioned, successfully resisting the encroachments of all other cells until the experimental program is modified.*

During any conditioning program then, we have an unstable group which at first increases in size even though gradually losing a few members, and a stable group which ultimately gains substantially complete control of the conditioned reflex. At any particular test trial the conditioned response is due to the combined efforts of both groups.

It is important to note that the conditioning process attracts *any* delta cell, active during the effective duration and not interfered with, but it is to the extinction process that we owe elimination of the unstable group. Furthermore, the process involves selective reinforcement on a basis of latency and gamma phase. Conventionally we can have reinforced trials, extinction, and selective reinforcement all in neat separate categories. In actual practice all are present during the course of a reinforced trial because we are dealing with an intact organism which contains the complete mechanism of intelligence.

The particular cells which form a stable group are determined entirely by their properties *and* the experimental program. Note that while it is the delta cell which becomes conditioned, the properties involved include those of the receptors, etc. Cells stable under one set of conditions may not be under another set.

The probability of conditioning for any delta cell is a function of pulse rate, among other things, and is not influenced by the factor of gamma phase which is an independent variable affecting stability. The amount of conditioning on any trial would always be a fixed percentage (probability) of the unappropriated motor cells, and a simple equation would result were it not for the fact that some of the unstable group give up their connections and all motor neurons do not have the same pulse rate.

The principal conclusion we can glean from the above verbal argument is that the unstable group ultimately disappears while the stable group receives a gradually decreasing increment with each successive trial and the resulting growth curve approximates well-known learning curves.

Extinction

The preceding discussions have already prepared the reader for all the elementary non-quantitative aspects of extinction. Suffice it to note here, by way of summary, actuation of the conditioned stimulus without reinforcement exposes the conditioned cells to a probability of extinction which results in a curve approximating the so-called law of organic decay.

Simple Differentiation

There are three types of differentiation, one of which is "pure," and two of which occur in conjunction with the first type.

The first and simplest type of differentiation is based on the gamma phase. While this type may not be conventionally included as differentiation it possesses the usual characteristic of transferring conditioning from one group of delta cells to another by means of selective reinforcement.

Assume a homogeneous field of sensory neurons with the usual variations

of latency, "threshold," and gamma phase, and assume further that a habit has been established to a particular experimental program involving no non-reinforced trials. The differentiation is effected by occasionally, and at random intervals, omitting reinforcement. The effect of this procedure is to re-define the stable and unstable groups with the result that the conditioning is gradually transferred to the group possessing a long enough gamma phase to survive the process undisturbed. Experimentally, the animal becomes immune to occasional non-reinforced trials.

The process can be continued to include delta cells having longer and longer gamma phases, increasing the number of successive non-reinforced trials, until terminated finally by an insufficient population of delta cells of long enough gamma phase to appropriate all the involved motor cells. When the appropriation fails to be complete the differentiation changes character to the extent that it is no longer possible for it to become and remain substantially independent of the relative number of reinforced and non-reinforced trials.

The second type of differentiation is based on latency. Obviously, if an established conditioned reflex is altered by increasing (or decreasing) the delay of reinforcement, we re-define the stable group on a basis of latency and after sufficient practice a test trial will reveal the change. This process can be continued with increasing delay or with trace reflexes until terminated by an insufficient population of delta cells possessing a long enough latency. Not so obvious, however, is the effect of the simultaneous and unavoidable gamma-phase differentiation which takes place. When we delay reinforcement to establish differentiation on a latency basis, the group of cells which bridge the gap of stage one receive an "increment of immunity" to non-reinforced trials. Although the gamma function cannot discriminate between non-reinforced and delayed trials and tends to resist a change of increasing delay, the distribution of the gamma and latency variables favors the latter.

The third type of differentiation is based upon receptor differences and, though occurring in a great variety of forms, involves no new principles beyond the use of multipolar neurons.

It will be observed that *any* difference in sensory neurons may become the basis of differentiation if the inter-relationships among the neurons and the relationship to the experimental program are appropriate. The three types of differentiation provide a mechanism for the correlation of large tracts of data: some of these applications are the subject of the following sections.

APPLICATIONS

We begin at last the interesting task of interpreting familiar examples of behavior in terms of the postulated mechanism.

The applications which follow are based upon phenomena selected for range of interest and generality. Necessarily omitted for lack of time and space are numerous other aspects of the BA which were investigated during the course of this research. To some extent the discussions become successively more abbreviated in certain respects to prevent excessive repetition and as a result the reader may find it advantageous to read the various applications in the order presented even though there is no serial functional relationship.

External Inhibition

External inhibition is the term applied, classically, to the momentary diminution of conditioned response when caused apparently by the intrusion into the experimental program of

an "extra" stimulus approximately in conjunction with the previously-established conditioned stimulus (5, p. 44). Since the classical concept of inhibition, as an active suppression of conditioned response, is not part of the BA theory, an explanation is to be sought in terms of some other mechanism. The *E* function seems adequate to account for the facts though at present no quantitative results are available. It is also necessary to postulate an orientation reflex in order to complete the explanation. For the present we assume that the orientation reflex excites various kinesthetic receptors which, by actuating their delta cells, depress *E*.

If we take the point of view that the conditioned stimulus is actually a signal, it naturally follows that the strength of the habit and its overt response should be related to the strength of the unconditioned stimulus, not to the magnitude of the signal which evokes it. The *E* function (Postulate 15) is designed for this property by decreasing *E* to compensate for increase in pulse rate of signal and thereby causing pulse rate of responding to be relatively constant. Strong signals give high pulse rates which are approximately offset in the motor cell by a corresponding reduction in *E* (Postulate 14). If now an extra stimulus is applied it also reduces *E* but without contributing a corresponding *conditioned* output pulse rate, hence, the conditioned response is lowered in magnitude, or disappears if the disturbing stimulus is sufficiently strong and effectively timed. This much requires no subsidiary assumptions. In addition, and oftentimes more important, the extra stimulus excites the orientation reflex, depressing *E* still further. Of course, if the extra stimulus is applied regularly the orientation reflex declines in strength while *E* tends to recover.

Referring to Fig. 10 we see there are two factors of significance: amount

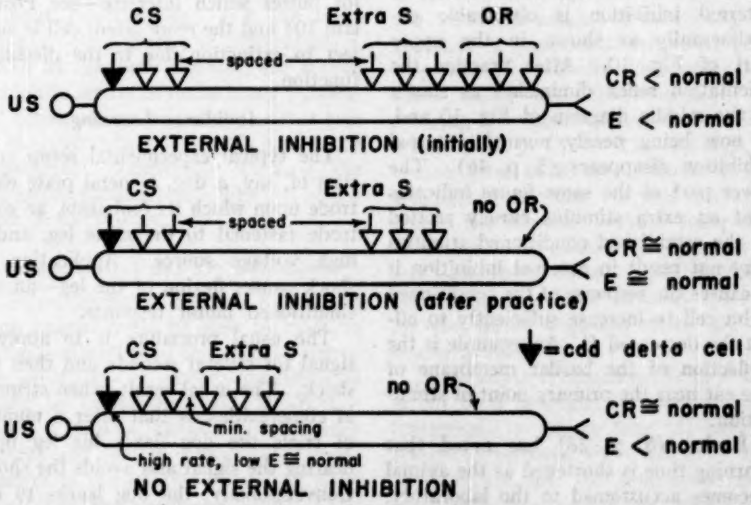

FIG. 10. Each motor cell above indicates a separate status of external inhibition. No significance attaches to the serial relationship of the delta cells to the axis of the motor cell. (See text for discussion.)

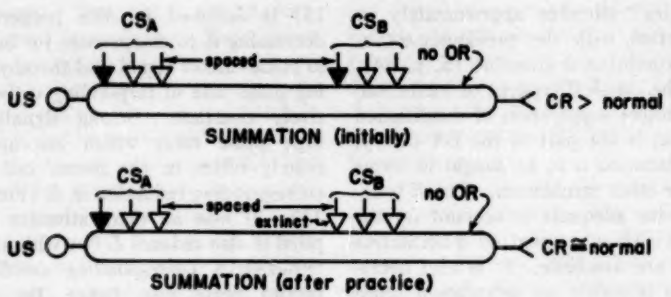

FIG. 11. The physical separation of the two delta cell groups results in a greater than normal *E* and consequent increased CR.

of practice with the extra stimulus and proximity on the motor cell body of the two groups of delta cells. Assuming that the extra stimulus is not a close relative of the established conditioned stimulus and that this relationship is maintained on the motor cell body we see at once that the extra stimulus or its derivative, the orientation reflex, must be fairly strong in order to establish a noticeable *E* gradient in the motor cell body. With this initial condition external inhibition is observable experimentally as shown in the upper part of Fig. 10. After practice the orientation reflex diminishes as shown in the middle diagram of Fig. 10 and, *E* now being nearly normal, external inhibition disappears (5, p. 46). The lower part of the same figure indicates that an extra stimulus closely related to the established conditioned stimulus may not result in external inhibition if it causes the response of the conditioned delta cell to increase sufficiently to offset the decreased *E*. An example is the deflection of the basilar membrane of the ear near the primary point of stimulation.

Pavlov (5, p. 28) has noted that learning time is shortened as the animal becomes accustomed to the laboratory. The same type of mechanism applies except that it is the *Q* function which is reduced in magnitude by extra stimuli. As the orientation reflex declines in strength the rate of learning increases.

The arpeggio paradox of Humphrey (4) is attributed to the same mechanism as external inhibition.

Summation

Figure 11 depicts the phenomenon of summation of conditioned reflexes (6, p. 313). The pulse rates add (except for pulses which interfere—see Postulate 10) and the more latent cell is subject to extinction due to the disabling function.

Avoidance Learning

The typical experimental setup consists of, say, a dog, a metal plate electrode upon which its foot rests, an electrode fastened to the same leg, and a high voltage source. Application of shock causes flexion of the leg—an unconditioned motor response.

The usual procedure is to apply a signal for several seconds and then the shock. The usual result, when stripped of complexities, is that after a number of trials the dog flexes his leg upon hearing the signal and avoids the shock. Conventionally, the dog learns to flex its leg and is rewarded by avoidance of expected shock; accordingly the habit should persist indefinitely since it is

regularly reinforced, yet it is generally conceded that the dog must be "reminded" now and then (2, p. 94).

The BA theory indicates that extinction starts with the first successful conditioned response because the shock, which is the reinforcement, is not applied but is avoided. After several such non-reinforced trials the weak habit has declined in strength to the point where it no longer suffices to cause flexion within the signal period and reinforcement is again automatically applied. The decline in strength has been selective since, as usual, the delta cells of shorter gamma phase and high decrement factor have contributed most to the loss. Conditioned cells with a sufficiently long gamma phase have been untouched by the chain of events. The habit strength builds up again but this time with fewer trials because of the incomplete extinction (it is assumed that an overt response requires some finite minimum amount or percentage of conditioning). Reaching sufficient strength the habit is again actuated without reinforcement and the cycle repeats. By a process of differentiation, based on the gamma phase, the animal gradually learns to make many successful responses in series but it never gets to the point where it doesn't have to be "reminded" occasionally because the maximum gamma phase duration sets a limit to the process.

Generalization and Differentiation

Generalization is essential to survival —an animal must be able to react to an approximation of the original stimulus situation or useful conditioning would be impossible; and, differentiation if it is to be "durable" must finally become and remain substantially independent of the relative number of reinforced and non-reinforced trials.

Generalized neurons are provided for in Postulate 6. For this section we are assuming that a generalized neuron has a receptor arrangement such that it responds when any neuron in the same field is stimulated. The exact properties are immaterial for the present discussion provided only that the postulated mechanism is possible. The degree of reaction is assumed independent of location of stimulation within the field—it is not only generalized, it is equalized.

In living intact organisms adjacent receptors are functionally related merely because they are physically imbedded in the same tissue. When the auditory mechanism, for example, is stimulated with a pure sine wave of a given frequency the basilar membrane deflects and receptors of adjacent frequencies are also stimulated but to a lesser degree (1, p. 44). This in itself is generalization but probably of insufficient degree to account for the facts. If now we invoke the participation of the above-mentioned generalized neurons and superimpose the peaked localized action on the equalized action we get the observed results: conditioning on the basis of a particular frequency gives responses of gradually diminishing size when tested with frequencies of increasing divergence from the original stimulus (3, p. 136). This assumes, of course, that both the specialized and generalized neurons compete for control of the response without being able to exclude one another. *They are in fact components of a compound stimulus as viewed by the nervous system; for the intact organism there are no simple stimuli.*

If the above arrangement is now subjected to a program of selective reinforcement, in which tone *A*, say, is always reinforced while adjacent frequencies are not, the generalized neurons and the unfavored specialized neurons enter an unstable group and surrender their control to tone *A* (5,

p. 120). As noted earlier, once a delta cell in the stable group appropriates a motor neuron, it successfully resists the encroachments of all other delta cells—differentiation becomes and remains substantially independent of the relative number of reinforced and non-reinforced trials.

Compound Stimuli

The individual components of compound conditioned stimuli appropriate a share of the available motor neurons roughly in proportion to their relative magnitudes (5, p. 142). This is true because the components are parts of a stable group, they compete for control of each motor neuron, and individually, they are unable to gain exclusive control of all the motor neurons.

If a compound of two components, A, the smaller component, and B, is entered into an experimental program in which the compound AB is reinforced and the component B when applied singly is never reinforced, we find that B loses its share of the conditioning to A and AB. Since periodic reconditioning is a factor, through the mechanism of gamma-phase differentiation, in this and many similar problems, Fig. 12 is set up to show in greater detail the process of transferring the conditioning to the other components—the weaker and the compound. It will be noted that the diagram shows individual delta cells on one motor cell but each delta cell represents a group when all motor cells are considered; also, the higher subscripts have been omitted.

Neurons B_1 and G_1 are driven to extinction while B_2 and G_2, possessing a longer gamma phase, easily survive occasional non-reinforced trials. Neurons B_2 and G_2 compete with A_1, A_1B_1, A_2, and A_2B_2 for control of the conditioned reflex and there is no way for one to displace the other. Therefore periodic reconditioning will, to some extent, prevent successful differentiation if the non-reinforced trials are limited to one at any time; in order for differentiation to components alone to be effective it is necessary to remove gradually the effects of periodic reconditioning by placing an increasing number of non-reinforced trials in series. This discriminates against periodic reconditioning and allows the latter group to gain nearly complete control. Of course if

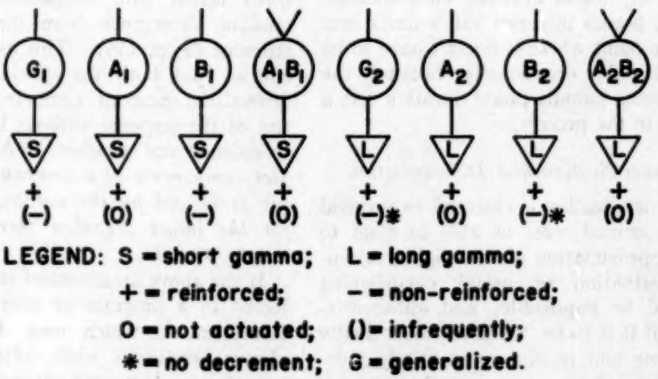

FIG. 12. The essential elements concerned in the establishment of classical conditioning, based on the bipolar component of a compound stimulus, are shown with the interfering elements of periodic reconditioning. The motor cell has been omitted. The same legend applies to subsequent figures.

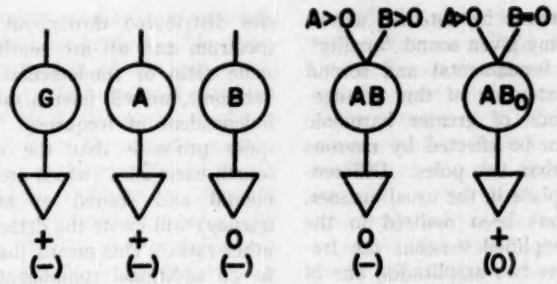

FIG. 13. Conditioned inhibition is here attributed to the so-called zero-ratio bipolar neuron.

the control is gained at a faster *rate* than periodic reconditioning the above effect will be obscured and the results will comply with the observations of Pavlov. Note that, in living organisms, A_1B_1 is probably a large component since one element is actuated by a relatively strong stimulus. On the other hand components with higher subscripts are relatively small as indicated by the distribution of decrement factor (Postulate 17). These facts favor successful differentiation in spite of interference by periodic reconditioning.

Conditioned Inhibition

Classically, conditioned inhibition is conceived to be an active suppression of conditioned response (5, p. 75). The evidence is impressive yet capable of other interpretations. The term *inhibition* has been applied to many phenomena now believed to be quite diverse in mechanism.

Let us assume that response is expected from application of tone *A* but not when accompanied by tone *B* and arrange reinforcement accordingly as shown in Fig. 13. The AB_0 connection is actuated only when *A* is given and *B* is zero, and, being regularly reinforced, is in the stable group which gradually dominates the conditioned reflex by appropriating all the available motor cells and excluding the generalized and *A* groups. It should also be noted that the "zero" pole is in effect an active inhibitory mechanism in the classical sense except for the fact that it is not, in itself, a learning mechanism since no conditioning takes place.

Stimulus Patterns

The BA approach to this problem is based on the properties of multipolar neurons. The author proposes to demonstrate that some of the complexities of auditory phenomena can be reduced to explanations in terms of mechanisms possessing relatively simple principles.

It is a matter of casual experience that we can recognize people we know by their voices alone. The differentiation, as usual, is based upon the environmental factor which remains constant throughout a series of trials. In this case it is the *relative* strengths of the harmonics in the sound. Figure 14 depicts the elements involved in the differentiation.

We have a generalized group *G*, the fundamental tone *F*, and two amplitudes of its second harmonic, H_1 and H_2. The bipolar neurons are equipped with a receptor sensitive to the fundamental frequency and one sensitive to the second harmonic, but the combining element, by assumption of particular properties, will relay the signal only when a specific ratio exists between the strength of stimulation of *F* and *H*. As a result of this provision the two bipolar

neurons will never be actuated simultaneously by any given sound "quality" consisting of fundamental and second harmonic. Extension of this arrangement to sounds of greater harmonic complexity can be effected by neurons having more than two poles. Differentiation takes place in the usual manner.

One fact has been omitted in the above over-simplified version: the frequency, *H*, has two amplitudes, one of which is reinforced and one is not. An increase in amplitude of stimulation actually brings additional receptors into action or, since the postulate of pulse rate includes no finite threshold, increases probability of conditioning to a significant level. Some of the monopolar neurons therefore are actuated on only part of the trials and differentiation is possible outside the specialized bipolar groups. Except in a laboratory, however, the intensity of sound varies considerably and consequently the bipolar mechanism provides the only stable group and hence controls the response.

Sounds, of course, are not confined to one fundamental frequency. It is possible, however, to place a multitude of bipolar combining elements in parallel with their outputs leading into one conducting element. If the bipolar elements have their fundamental frequencies distributed throughout the audio spectrum and all are sensitive to the same ratio of fundamental to second harmonic, we will have a ratio detector independent of frequency. Now it is quite probable that the second and fourth harmonics (which are the fundamental and second of another frequency) will excite the detector of some other ratio. This means that there will be an additional component competing for control and one which can be eliminated by selective reinforcement. The rest is familiar ground.

At this point we remember the existence of the latency elements and perceive at once that if we don't restrict ourselves to harmonic relationships, the problem of recognizing a melody in an unfamiliar key is automatically included in the above discussion.

Instrumental Conditioning

In the discussion on *avoidance learning*, which was not explicitly treated as instrumental conditioning, we assumed a noxious stimulus which would directly actuate the desired unconditioned response. This is equivalent to exclusion of spinal mechanisms; in fact, the BA theory, as delineated in this paper, is restricted almost entirely to the conditioning mechanism itself.

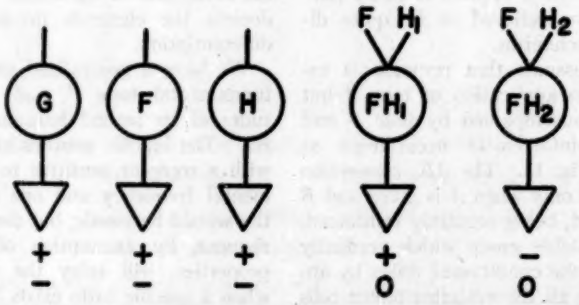

FIG. 14. Sound-pattern discrimination is but a form of differentiation identical in principle with previous examples.

Instrumental conditioning naturally involves spinal mechanisms. The structure and properties of the spinal cord lend support to the belief that any stimulus affects all final common paths to some extent even though the amount may not be easily detectable in the overt behavior. Furthermore, although some path is ordinarily prepotent for a particular stimulus, it is probable that the static hierarchy of responses is often violated by various random stimuli and the effect of varying thresholds. No doubt there are limitations to these processes but it is conceivable that differentiation operating on such a mechanism is sufficient to bring about one of the observed results: escape from a noxious stimulus by means of a specified, distantly related, response (escape learning).

We propose to demonstrate in this section only that if the noxious stimulus actuates several motor *nerves* in succession, if the hierarchy of responses is not perfect, and if the experimental program terminates the noxious stimulus upon elicitation of a specified response, differentiation is an adequate mechanism to explain escape learning.

Without denying a direct stimulus-response relationship we assume that the experimental environment is not completely controlled, that variations in posture exist, and that these are sufficient to cause variations in unconditioned responses even under seemingly ideal theoretical conditions. The situation, then, can be expressed in terms of probability: A chance relationship exists between the timing of the noxious stimulus and the various possible combinations of postural reflexes, thresholds, etc. The static hierarchy of responses can be assigned numerical values of probability but these values fluctuate from moment to moment with the result that a perceptible degree of randomness exists.

With the above assumptions all possible responses will be forthcoming in time and each response can be expected to fail occasionally. Each stimulus that elicits a response is reinforcement and causes conditioning between it and the conditioned stimulus which in this case must also be the shock. Since there is a hierarchy of probability of responses we expect the most probable to condition at the highest rate and expect overt escape learning to be delayed roughly in proportion to the degree of "remoteness" between noxious stimulus and specified response. Conditioning, and responding, act to alter the effective thresholds and hierarchy of responses so that "incorrect" learning is indicated until the conditioned specified response has sufficient strength to compete on more favorable terms. As the competition improves, the probability is greater that the specified response will be elicited before one or more of the conditioned incorrect responses, and each time this occurs the reinforcement is terminated in accordance with the experimental program and some decrement occurs in each *nerve* not reinforced (the conditioned stimulus is excited for all nerves).

We now have the familiar situation of stable and unstable groups—each trial continues until the specified response is elicited. But these groups are not competing for the same motor cells and they cannot directly exclude one another as with classical conditioning. Nevertheless the incorrect responses gradually decline in frequency due to the gradual elevation of the specified response to the head of the hierarchy. With this theory we never expect perfection in response but only a relatively high probability.

Since the shock is both conditioned and unconditioned stimulus we have to account for the fact that conditioning occurs in spite of their apparently

simultaneous application. It is believed that the unconditioned responses of the intact animal subjected to the conditioning routine are actually due to compound stimuli (the postural reflexes, etc.) of which the shock is but one component. If this belief is well founded it is apparent that each response is elicited separately as part of a more or less random sequence and hence all but the first must occur *after* the beginning of the conditioned stimulus. No obstacle to the theory from this source occurs until the stage is reached when the specified response is elicited first. Here we can make additional assumptions, or better yet, plead ignorance of the exact facts.

With hunger drive in place of shock the principle is substantially identical. The reader can reach this conclusion by assuming that hunger merely sensitizes the locomotion and postural system to the point where the usual random stimuli are effective unconditioned stimuli. It is also postulated, informally as yet, that food causes a momentary dynamic reduction in drive (or the animal would continue locomotion and leave the scene of the food) which terminates the effect of further reinforcing stimuli.

The conditioned stimuli for instrumental conditioning may be either simple or compound but only the compound stimuli possessing a component representing the drive are useful because they are the only ones which will be actuated when reinforcement is probable.

Consider finally the mechanism common to these phenomena. A dog flexes his leg *and* (not *to*) escapes shock (or hunger drive, etc.), successful escape maintains the habit at the top of the hierarchy, and failure to escape drives the habit to extinction *if, and only if, another response is successful.* Avoidance learning follows exactly the same pattern and is in fact escape learning up to the point where avoidance succeeds, when, reinforcement being nil, the habit declines and has to depend on periodic reconditioning.

SUMMARY

Although we have endeavored at times to raise a superstructure where the foundation seemed inadequate, certain contributions toward the principles of intelligent mechanisms may have enduring value. Reference is made to the demonstration of the necessity for stable stimulus differentiation and the fact that a functional mechanistic principle has been provided. Likewise, the multipolar neuron, which is not a structurally accurate representation of anatomical data, nevertheless presents a functional principle of more than transitory worth. But most important of all, and in the words of the first sentence of this paper, the Brain Analogy is an instrument for research; it is expected that a physical basis for behavior theory will accelerate progress.

REFERENCES

1. GALAMBOS, R., & DAVIS, H. The response of single auditory-nerve fibers to acoustic stimulation. *J. Neurophysiol.*, 1943, **6**, 39–57.
2. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning.* New York: D. Appleton-Century Company, 1940.
3. HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125–148.
4. HUMPHREY, G. The effect of sequences of indifferent stimuli on a reaction of the conditioned response type. *J. abnorm. soc. Psychol.*, 1928, **22**, 194–212.
5. PAVLOV, I. P. *Conditioned reflexes.* (Trans. by G. V. Anrep.) London: Oxford University Press, 1927.
6. RAZRAN, G. H. S. Studies in configural conditioning: I. Historical and preliminary experimentation. *J. gen. Psychol.*, 1939, **21**, 307–330.
7. SKINNER, B. F. *The behavior of organisms.* New York: D. Appleton-Century Company, 1938.

[MS. received June 30, 1950]

TWO KINDS OF EXPERIMENT DISTINGUISHED IN TERMS OF STATISTICAL OPERATIONS

BY MELVIN R. MARKS

Personnel Research Section, AGO

The work which a scientist does when he is "finding out" something has been variously called research, investigation, study, inquiry, examination, etc. Underwood (9) dichotomizes experimental problems into the "I-wonder-what-would-happen," and the "I-bet-this-would-happen" types. Mann [1] selects for distinction the terms "inquiry" and "research": The former proceeds without theory, whereas the latter is generated by theory and is reflexive, *i.e.*, the results of research modify the constructs of the basic theoretical matrix. Both Underwood and Mann point out that basically there is a difference between just "finding-out," and "finding-out-if-a-hypothesis-is-consistent-with-fact." From this standpoint a distinction might be made on the basis of necessity for control. In "finding-out" control might be dispensed with, since control must always be with respect to something, and in "finding-out" that something is what is sought. On the contrary, in "finding-out-if- . . . ," control is necessary properly to evaluate results.

The writer believes that some sort of distinction is necessary, that it is best made in terms of statistical operations, and that it has important implications for the validity of conclusions drawn from results. What is needed first are two terms relatively free of connotations bearing on possible distinctions. While "inquiry" has been little used, "research" has been overused—and in a wide variety of meanings. To avoid ambiguity, the terms "experimentation$_{I}$," and "experimentation$_{II}$" have been selected.

A brief review of the statistical concepts relevant to the distinction will be helpful. Most current texts discuss them in part, but it is believed that the synthesis proposed here is new. At any rate such synthesis should be of value to those psychologists engaged in "experimentation" who have not found time to peruse thoroughly the introductory material of the more advanced statistical texts.

Statistical Hypothesis. This is any testable assumption about parametric values. Typical parameters are population mean, population variance, difference between two population means, etc. As examples, we might make several statistical hypotheses about the populations from which two sample groups, say A and B, were drawn:

"There is no difference between the means of the populations from which A and B were drawn"—algebraically,

$$(m_A - m_B) - 0 = 0. \qquad (1)$$

"The population means of A and B differ by 5 points."

$$(m_A - m_B) - 5 = 0, \qquad (2)$$

or, more generally,

$$(m_A - m_B) - K = 0$$
$$(K = \text{any constant}). \qquad (3)$$

Note that in each of the three equations, the quantity $(m_A - m_B)$ is given a positive sign. In effect, a definite prediction is being made about the direction of the difference in magnitudes. A completely generalized equation,

$$(\ |m_A - m_B|\ ) - K = 0, \qquad (4)$$

[1] From unpublished seminar discussions.

states, "the absolute difference between the means of the populations from which A and B were drawn does not differ from some constant K."

It is assumed frequently that (4) represents the classical "null hypothesis." Actually, (1), (2), and (3) are null hypotheses as well. In each case the hypothesis is that the parametric difference between some assumed difference and a given constant is zero. In each instance the right side of the equation is the null which is to be tested.

Level of Confidence, Statistical Significance, etc. When the test of a statistical hypothesis yields an estimate of the probability that the null is tenable, it is customary to say that the results are "significant" if the associated probability, p, is less than some value chosen arbitrarily in advance of the test. For the purpose of this paper the statement, "significant at the 5 per cent level," means that the hypothesis tested was of the form (1), (2), (3), or (4), and that the associated $p < .05$.

Type I Error. Suppose that A and B are each groups of 11 children for whom I.Q.s have been determined. Let the mean difference in I.Q. be 4.172, and the standard error of that difference be 2. With these data, $t = 2.086$, $d.f. = 20$, and $p = .05$. An investigator using the 5 per cent level would conclude that there is a significant difference. Two possibilities exist: either there is a difference in the population means, or there is not. If, actually, there is a difference, his conclusion would be correct; if, actually, there is no difference, *i.e.*, both groups come from the same population, he has erroneously falsified the null. The false rejection of the null hypothesis is called a Type I Error. Note that the probability of occurrence of a Type I Error is determined absolutely by the level of confidence adopted. In the example the 5 per cent level was chosen; hence, by definition, there are 5 chances in 100 of a random fluctuation which would lead to a t significant at that level.

Type II Error. This may be illustrated by recasting our example. Suppose that the two groups of children came from populations whose means differed by 2 I.Q. points; that the difference between the means of the samples was also 2 points (*i.e.*, "perfect" sampling), and that the standard error estimated from the sample difference was again 2 points. In this case, $t = 1$, $d.f. = 20$, $p > .30$. The same investigator would conclude erroneously that the null was tenable. The failure to reject the false null is called a Type II Error.

The *power* of a statistical test is defined as the probability that a Type II Error will *not* occur. Unlike the case of the Type I Error, the probability of a Type II Error cannot be fixed by the investigator, since it is a function of three variables, only two of which are under the investigator's control. The variables are: (1) the sample size; as N increases probability of Type II Error decreases (power increases), *i.e.*, a parametric difference is more likely to be found if the sample is large; (2) the probability of Type I Error (per cent level of confidence chosen by the investigator); as the level increases (in the sense that a 5 per cent level is an increase over the 1 per cent level), probability of Type II Error decreases (power increases), *i.e.*, it is more likely that a parametric difference will be detected with the 5 per cent level than with the 1 per cent level; (3) the relative magnitude of the parametric difference; as the magnitude (in standard error units) increases, probability of Type II Error decreases (power increases), *i.e.*, if the population difference is large it is likely that such difference will be detected in the sampling.

It is obvious that (3) above is unknown to the investigator. It should be noticed that, unless the hypothesis tested includes a definitive statement of the relative magnitude of the parametric difference, the power of the test is indeterminate. For example, given the hypothesis that $m_A > m_B$, without qualification as to the magnitude of the difference, the power cannot be computed; but, if m_A and m_B are assigned definite values, the power may be computed. Neyman and Tokarska (7) have tabulated the power of the *t*-test for various levels of confidence (Type I Error), sample sizes, and parametric differences. For example, with 20 *d.f.* at the 5 per cent level, if the parametric difference is 4.12 standard error units, power = .99, *i.e.*, probability of Type II Error = .01; with 1.70 parametric difference, power decreases to .50; with .39 units of parametric difference, power decreases to .10. Increase in power with increase in *N* is slight. For infinite *d.f.* at the 5 per cent level, 3.97 units of parametric difference give power of .99; 1.64 give power of .50, and .36 give power of .10. Johnson (5, p. 68) cites several papers where power functions of other tests are tabulated.

Two-Tailed Tests. In the two examples concerning the difference in group I.Q., the *sign* of that difference was not stipulated. Those examples were tests of the null expressed as in equation (4). The sign of *t* was not at issue; m_A could have exceeded m_B or vice versa. Tests made on this basis are called two-tailed tests, because both tails of the symmetrical distribution of *t* or x/σ are considered in determining the associated probability. The tables of *t* and x/σ are misleading in that only positive values of the statistic are tabulated, but the associated probabilities are for two-tailed tests. Although the distribution of Chi^2 is not symmetrical, here too tabular probabilities are for two-tailed tests, *i.e.*, the probability of frequency discordance in *either* direction from the hypothetical; note, however, that Chi^2, being a sum of squares, is necessarily always positive. Similarly, the *F*-ratio is always positive, but here too, tabled values refer to probabilities of chance *plus or minus* differences among the means of the treatments—thus the *F*-test as usually used is a two-tailed test.

One-Tailed Tests. Suppose that, in the example concerning the children's I.Q.s the investigator had predicted *before* the measurements were taken that $m_A > m_B$. The experimental hypothesis would fail obviously if $m_A < m_B$, by no matter how slight an amount. The investigator in making his test should exclude exactly half of the values which *t* might have, since he is interested only in the occurrence of *t*'s which bear a positive sign. This is quite legitimate since, in effect, he has agreed that all negative *t*'s will indicate an insignificant difference. His question is, given that m_A does exceed m_B, is the difference too great to be attributed to chance variation at the level of confidence selected. The two-tailed test does not apply here, since only positive values of *t* are considered. The hypothesis is properly evaluated by using only one tail of the sampling distribution. The *t* sought is that which might occur 10 per cent of the time if the null were true, since 5 per cent, or half of the possible values, would occur in the positive direction. This $t = \pm 1.725$, since $+ 1.725$ would occur by chance 5 per cent of the time. Recall that the 5 per cent value for the two-tailed test is 2.086, and it will be seen that the investigator is much more likely to detect the parametric difference if he uses the one-tailed test.

The one-tailed test leads to a similar increase in power with other statistical tests as well. For the sampling distribution of x/σ, ± 1.96 and ± 2.58 are

the values required for the 5 per cent and 1 per cent levels, respectively, when a two-tailed test is employed; for the one-tailed test the corresponding values are 1.65 and 2.33. When using the *Chi*^{2} and *F*-tables with a one-tailed test, the value of the statistic which has an associated probability *double* that of the level of confidence selected is the value which must be reached for falsification of the null. That value is always less than would have been required if a two-tailed test were used.

We are now prepared to examine, or better to define, the distinction between experimentation$_{I}$ and experimentation$_{II}$. *Experimentation$_{I}$ is that form of experiment to which two-tailed tests of the data are appropriate; experimentation$_{II}$ is that form of experiment to which one-tailed tests of the data are appropriate.*

It follows that (a) preëxperimental predictions and (b) use of one-tailed tests are severally necessary and together sufficient for experimentation$_{II}$. The *intention* of the experimenter is *not* the controlling factor. Thus, if he decides to do experimentation$_{II}$ but either fails to make a prediction or employs a two-tailed test he is actually engaged in experimentation$_{I}$. Conversely, if he decides on experimentation$_{I}$ but predicts the direction of his results and employs a one-tailed test, he is actually engaged in experimentation$_{II}$.

The experimental distinctions are not mere verbal quibbles. We may exemplify their importance first with a fictitious problem, then with illustrations chosen from psychological literature.

Suppose that the investigator of group differences in I.Q. found that his sample means differed by 3.9 I.Q. points, and that the standard error of the mean differences was again 2.0. On these data $t = 1.95$. If we suppose further that group *B* had been given special coaching designed to "increase" I.Q., while group *A* had been maintained as a control, it would appear that the investigator thought he was engaged in experimentation$_{II}$. However, unaware of the distinction made here, he consults the *t*-table and finds that the obtained value of 1.95 falls short of the 2.086 required for the 5 per cent level. He concludes that the special treatment given to group *B* was ineffective in increasing I.Q. Actually, by using the right test (one-tailed), he would fulfill the requirements for experimentation$_{II}$, and his results would be significant at the indicated level.

It is not difficult to find instances where investigators have done experimentation$_{I}$, under the illusion that they were doing experimentation$_{II}$. Tinker (8) had *Ss* read two forms of his Speed of Reading Test, the first under "ordinary" conditions, the second while the reading table was vibrated mechanically. He predicted that vibration would cause fatigue which would in turn lead to a decrement in reading speed. His results bore out the prediction. He also used a control group who read both forms under "ordinary" conditions. Here there was a critical ratio of 2.52 in favor of the first reading. He employed a two-tailed test at the 1 per cent level and found the difference could be attributed to chance (2.58 was needed). He might have predicted that there would be some fatigue in the second reading, even though the table was not being vibrated. Such predicted decrement with a one-tailed test would have been significant at the 1 per cent level. Tinker's results with the control group amount to experimentation$_{I}$, not experimentation$_{II}$.

Kircheimer, Axelrod and Hickerson (6) studied college students who: (1) changed their majors and (a) were counseled, (b) were not counseled; and (2) did not change their majors and (a) were counseled, (b) were not counseled. They predicted that, of the groups

who changed majors, those who were also counseled would show the greatest increase in grade point average. The critical ratio (in the direction of the prediction) was 1.84. The writers said that this was significant at the 7 per cent level, and had they employed a 5 per cent level criterion, the result would not have been significant statistically. Had they employed a one-tailed test as they should, the result would clearly have been significant at the 5 per cent level. These writers too have failed to draw the distinction between experimentation$_{I}$ and experimentation$_{II}$ and their results have appeared less critical thereby.

Gagné (4) predicted that, when highly similar items were grouped together, learning of serial lists would be accomplished in a shorter time than when such items were separated. The results were in the predicted direction with a critical ratio of 1.69. Gagné stated that the probability of occurrence of a deviation in the expected direction was 4.6 per cent. He correctly used a one-tailed test. Had he used the two-tailed test and adopted the 5 per cent level, he would not have found significant differences, and the prediction would appear to have failed. He rightly distinguished between experimentation$_{I}$ and experimentation$_{II}$.

The example concerned with Tinker's results was predicated on the basis that the direction of the decrement in reading decrement *could* have been predicted *before* the data were in. It must be emphasized that the one-tailed test is not justified unless the prediction is made prior to the data. If an investigator begins a study without preconceptions of results, and on studying those results generates a theory which will account for them, he cannot accept such afterthoughts as predictions, *i.e.*, switch —on the spot—from a two-tailed to a one-tailed test and pride himself on doing experimentation$_{II}$. He can *use* the results to predict *future* data. If he makes the switch after collecting the data, he has, in effect, increased the probability of committing a Type I Error without realizing it. If he uses the 5 per cent level on a one-tailed test, this is equivalent to the 10 per cent level with a two-tailed test, and he has unwittingly doubled his level of confidence: a dangerous departure from scientific conservatism. Burke (1) and Cronbach (2) have pointed out the inflation of level of confidence which occurs with either adding more subjects on the basis of already gathered data, or "shopping around" in compiled data for comparisons which are likely to show significant differences.

The experimenter$_{II}$ is in possession of a statistical tool not at the disposal of the experimenter$_{I}$. He is not limited to the prediction of results in a particular direction, but he may frame his hypothesis in such a manner that the magnitude of the parametric difference is included as well. For example, again referring to the comparison of group I.Q.s mentioned previously, the research worker might use an hypothesis of the form of equation (3), where the stipulated difference attributable to the "coaching" may be arbitrarily small. In this case, if the null remains tenable in light of the data, the theory is not rejected, whereas in experimentation$_{I}$, the null is of the form of equation (1) and, if there were theory, that theory would appear to have failed. This tool should be used with caution. No theory is particularly valuable if it predicts differences which are impracticably small. For instance, if the same coaching were expected to increase I.Q. by .5 points, the theory might remain tenable but the question "So what?" suggests itself. In this vein, Edwards (3) has distinguished "practical" from "statistical" significance. The experiment-

er_{II} should ever attempt to increase the power of his tests. If he is at liberty to set up experimental situations he should endeavor to maximize parametric differences since these contribute to test power. It follows that, in new areas $experimentation_{II}$ might wisely be limited to the comparison of extreme instances, rather than include the attempted distinction of subtle differences.

In summary, $experimentation_{II}$—as distinguished from $experimentation_{I}$—is characterized by predictions and one-tailed tests of significance. It has inherently greater statistical power and increases the productivity of investigation.

REFERENCES

1. BURKE, C. J. The effect of postponed decisions on statistical tests. Paper read at 22nd annual meeting of the Midwestern Psychological Association.
2. CRONBACH, L. J. Statistical methods applied to Rorschach scores: a review. *Psychol. Bull.*, 1949, **46**, 393–429.
3. EDWARDS, A. L. *Experimental design in psychological research*. New York: Rinehart, 1950.
4. GAGNÉ, R. M. The effect of sequence of presentation of similar items on the learning of paired associates. *J. exp. Psychol.*, 1950, **40**, 61–73.
6. KIRCHEIMER, B. A., AXELROD, D. W., & HICKERSON, G. X., JR. An objective evaluation of counseling. *J. appl. Psychol.*, 1949, **33**, 249–257.
5. JOHNSON, P. O. *Statistical methods in research*. New York: Prentice-Hall, 1949.
7. NEYMAN, J., & TOKARSKA, B. Errors of the second kind in testing Student's hypothesis. *J. Amer. statist. Assn.*, 1936, **31**, 318–326.
8. TINKER, M. A. Effect of vibration on reading. *Amer. J. Psychol.*, 1948, **51**, 386–390.
9. UNDERWOOD, B. J. *Experimental psychology*. New York: Appleton-Century-Crofts, 1949.

[MS. received July 26, 1950]

RESPONSE-SELECTION IN DISCRIMINATIVE LEARNING

BY PHILLIP WEISE AND M. E. BITTERMAN

The University of Texas

In a recent issue of this JOURNAL (13) Nissen contrasts two sets of terms which have been used to describe the behavior of animals in discrimination-situations. One set of terms is response-oriented; the animal is reported to "turn to the right" or to "turn to the left" in the presence of a given pattern of stimuli. This description implies that each spatial arrangement of the stimuli to be discriminated constitutes a discrete configuration for which the animal must select an appropriate response. The second set of terms emphasizes the acquisition of functional significance by afferent components and ignores the precise character of the movements involved in the solution of the problem; the animal is reported to "approach" one stimulus and to "avoid" the other. As Nissen remarks, "Even introspective evidence suggests a critical difference: 'reaching towards the right' (when green happens to be at the right, red at the left), is not the same thing as 'approaching (or reaching towards) green' " (13, p. 121).

Nissen recognizes that, for higher organisms at least, both terminologies may find application (according to the nature of the discrimination-problem in question), but he is inclined to believe that the approach-avoidance formulation will cover all cases of discriminative learning. Nissen describes an experiment with chimpanzees which may seem to suggest both the validity of the distinction which he emphasizes and the interpretation which he defends. A group of animals was trained to obtain food in an apparatus which presented black and white cards in an up-down relation (one above the other). The task of the animals was to select one of the two cards on each trial irrespective of whether it appeared in the upper or lower position. Other animals were trained in the same manner except that the cards appeared only in a right-left relation. When the animals of each group had reached the criterion of learning, they were shifted to the problem of the other group. That is to say, animals trained in the right-left situation were shifted to up-down, and conversely. Computations by the method of savings revealed a large amount of transfer from problem to problem (66 to 100 per cent), and Nissen concludes that these results can be accounted for only in approach-avoidance terms.

It must be noted, however, that Nissen's results are not unambiguous. Certainly one cannot dispute his interpretation of the behavior of animals showing 100 per cent transfer. Clearly these animals had learned something which was relatively independent of motor functions. As Nissen suggests, perfect transfer can only be explained if we schematize the learning as

$$S_{black} — R_{approach} \qquad S_{white} — R_{avoidance}$$

rather than as

$$S_{black\text{-}white} — R_{left} \qquad S_{white\text{-}black} — R_{right}.$$

But what can be said of animals showing only intermediate amounts of transfer? Some disturbance was produced either by the changed perceptual configuration, *qua* percept, or by the change

in the appropriate response, or both. Instances of less than complete transfer deprive the approach-avoidance formulation of complete generality.

The most explicit statement of the response-oriented interpretation of discriminative learning is to be found in the position of Gulliksen and Wolfle (5, 6). These writers maintain that each spatial arrangement of stimuli in a two-choice problem (*e.g.*, black-left/white-right and white-left/black-right) must be considered as a discrete configuration, and that the process of learning such a problem is essentially a process of selecting the response appropriate to each configuration. Although the evidence which Gulliksen and Wolfle present in support of this interpretation is not crucial, a recent experiment by Bitterman and Coate (2) has incidentally provided supporting data. In learning a brightness-discrimination on the black and white alleys of an elevated maze, the rats of Bitterman and Coate showed a significant tendency to master each spatial arrangement of the alleys at a different rate depending upon the nature and frequency of their experience with each situation in an earlier stage of training. If, as Nissen suggests, the animal learns only to approach one of the two stimuli and avoid the other, their spatial relation should make no difference—a deduction which is, in fact, the basis of his own experiment.

The position of Gulliksen and Wolfle has been severely criticized by Lashley (7) who presents convincing evidence for its lack of complete generality. Lashley trained rats in a three-choice discrimination-situation, with three circles differing in size as stimuli. His animals could master the problem when the largest circle was positive or when the smallest circle was positive, but not when the intermediate circle was positive. Since all three problems involve the same six spatial arrangements of the stimulus-cards and the same three responses (jumping to left, to right, or to center), the only difference being in the pairings of stimulus-arrangement and response, the view of Gulliksen and Wolfle does not make it possible to account for the failure of the animals to master the intermediate-size problem.

We have noted thus far that both formulations under examination are confronted with certain difficulties. Lashley's results with the three-circle problem cast doubt upon the general validity of the Gulliksen-Wolfle position,[1] and the approach-avoidance position, as formulated by Nissen, cannot deal with the evidence of situational learning provided by Bitterman and Coate, or with instances of less-than-complete spatial transposition reported by Nissen himself. But while Lashley's evidence is crucial, and undeniably limits the applicability of the theory of Gulliksen and Wolfle, Spence's elaborate and ingenious version of the approach-avoidance position (15) is not contradicted by the data considered up to this point.

Spence's theory falls rather clearly into Nissen's approach-avoidance category. As Spence has written: "Discrimination learning does not consist . . . in the strengthening of one response relatively to another or others as in the case of problem-box learning, but involves, rather, the relative strengthening of the excitatory tendency of a certain component of the stimulus complex as compared with that of certain other elements until it attains sufficient strength to determine the response" (15, pp. 429–430). The central assumption of Spence's theory is that discrimination is the end-product of a continuous, cumulative process of ac-

[1] There is other evidence which contradicts the position of Gulliksen and Wolfle. Lashley's succinct criticism serves only as an example of the sort of difficulty with which the theory is faced.

quisition of excitatory and inhibitory characteristics by components of the stimulus-situation. Consider, for example, a black-white problem which requires the animal to go to white irrespective of its lateral position. The course of discrimination may be described in terms of four principal afferent components, two of which are relevant (blackness and whiteness) and two of which are irrelevant (rightness and leftness). These components may have certain excitatory and inhibitory values even before the animal makes the first response (presumably derived from its previous experience). Suppose that the first trial is to black-left/white-right. If the combined excitatory values of blackness and leftness minus their combined inhibitory values exceeds the combined excitatory values of whiteness and rightness minus their combined inhibitory values, the animal will go to black-left (non-reinforced), thus increasing the inhibitory values of blackness and leftness. As the training continues (under conditions of spatial alternation) the irrelevant components are inconsistently reinforced, whiteness is consistently reinforced, and blackness is never reinforced, until the black-white difference exceeds whatever difference may exist between irrelevant components, at which time the problem is solved.

This theory can readily account for certain instances of situational learning in discrimination-problems. If, as in the experiment of Bitterman and Coate, the animal comes to the black-white problem with a position bias (a marked difference in the excitatory values of leftness and rightness), the spatial arrangement which presents the positive stimulus on the preferred side will be mastered more readily than the reverse arrangement. Spence's theory would also lead us to expect cases of less-than-perfect transfer in Nissen's experiment. Preëxisting differences in the excitatory values of rightness and leftness might cause difficulty for an animal shifted from the up-down problem to the right-left problem, and conversely.

Although most of the experimental work on discriminative learning has been done with problems of the *simultaneous* kind, in which two different stimuli are presented contiguously and the animal is required to choose one of the pair irrespective of spatial arrangement, another method is available which has crucial bearing on the problem at hand. Suppose that the stimuli to be discriminated are presented *successively*. For example, in a T-maze discrimination-apparatus the animal may be required to turn to the right when a tone of one pitch is sounded and to the left when a tone of a second pitch is sounded (12); or, the animal may be required to turn to the right when two white alleys are presented and to the left when two black alleys are presented (9, 10). Nissen recognizes that the ability of animals to master such problems poses a difficulty for approach-avoidance theory, which he attempts to save by assuming that such learning involves a complex, conditional type of solution. "The fact that such problems are in general learned with relative difficulty," he remarks, "is in keeping with this interpretation" (p. 129). Actually, however, there has been no systematic study of the relative difficulty of simultaneous and successive discrimination bearing on the theoretical question under consideration.[2] Such a study would pro-

[2] Grice (4) has used the terms "simultaneous" and "successive" in a somewhat different sense. His experiment was designed to study the effects of opportunity for direct comparison of the stimuli to be discriminated. Both of the problems mastered by his animals can be dealt with in simple approach-avoidance terms.

vide a crucial test of the approach-avoidance formulation.

The experiment to be reported was designed to compare the relative difficulty of simultaneous and successive light-dark discrimination-problems. At each choice-point in a four-unit alley maze, one group of animals encountered one lighted alley and one darkened alley. The task of these animals was to go either to dark or to light irrespective of lateral position. A second group of animals found either two lighted alleys or two darkened alleys at each choice-point and were required to turn in one direction in the first case and in the opposite direction in the second. Nissen has already predicted that the second problem should be more difficult, and Spence's position leads to the same prediction. If stimulus-components acquire functional properties *independently* and function together only in a *summative* manner, an animal confronted with the second problem must be expected, on any trial, to approach the preferred side (irrespective of the brightnesses presented at each choice-point) if there is a position preference, and either side at random if there is not. Learning cannot take place because there is no stimulus-component which is consistently reinforced or non-reinforced.[3] Lashley must also predict that the simultaneous problem should be easier, not only because it does not require a conditional solution but also because it provides the animal with opportunity for direct comparison of the stimuli to be discriminated (8). Only a theory such as that of Gulliksen and Wolfle leads to the prediction that the successive problem will be less difficult. The two configurations presented in the simultaneous problem (light-left/dark-right and dark-left/light-right) are probably more similar to each other and therefore should be more difficult to distinguish than those presented by

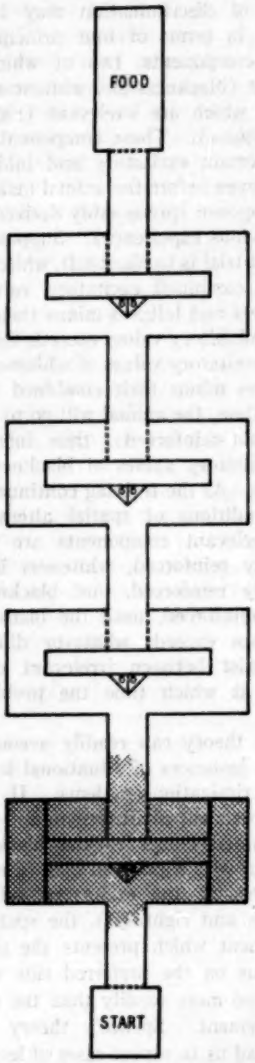

FIG. 1. Diagram of the apparatus.

[3] Mastery of the problem can be accounted for in terms of afferent neural interaction, but the extensive training required for stimulus-compounding leads to the prediction that the simultaneous discrimination will develop at a more rapid rate than the successive.

the successive problem (double-light and double-dark). The results of the experiment to be reported are in accord with this prediction.

APPARATUS AND PROCEDURE

The apparatus employed, which is illustrated in Fig. 1, was a modification of Stone's multiple discrimination-box (16). It consisted of a starting-box (8″ × 8″), a food-box (8″ × 17″), and four identically constructed discrimination-units, each incorporating a separate choice-point. Each alley was 12″ long and 4½″ wide. All walls were 4½″ high. The entire apparatus was painted flat black and the top was uniformly covered with heavy-gauge hardware cloth. The apparatus had no floor, but was set on a sheet of black oilcloth which could be cleaned easily.

The essential details of construction are indicated in Fig. 1. The hatched lines represent the hardware cloth top. The solid double lines drawn only in the first unit represent black metal bars used to support the sides. The double broken lines drawn at the starting-box and food-box show the positions of guillotine doors operated by the experimenter from a position just back of the starting-box by means of a string and pulley system. The double broken lines in each discrimination-unit indicate the points at which partitions could be inserted in order to convert either alley into a cul. Each choice-point contained two 7½ w., 115 v. lamps separated by a white partition which functioned as a reflector. The lamps were set in a V-shaped housing of clear plate-glass. Diffuse, indirect light from a fixture 6′ above the center of the maze dimly illuminated all of the alleys, thus permitting the animals to see the interior of the apparatus. Nevertheless, when one of the two lamps at a given choice-point was on and the other off, a sharp difference in the brightness of the alternative pathways was produced. The experiment was conducted in the Dome Room of the animal laboratory at the University of Texas.

Twenty naïve albino rats ranging in age from 85 to 100 days were employed. The animals were divided into two groups. Group I, consisting of six males and four females, learned the simultaneous problem, and Group II, five males and five females, learned the successive problem.

In the simultaneous problem, one lamp at each choice-point was on and the other was off. The task of the animals was to turn either in the direction of the lighted alley or of the darkened alley at every choice-point. Half the animals of Group I were required to go to light and half were required to go to dark. An incorrect turn at any choice-point led to a barrier, while a correct turn led to subsequent choice-points and eventually to the food-box. On each trial the stimuli were so arranged as to require two right turns and two left turns. There were six possible turning orders (RRLL, RLRL, RLLR, LLRR, LRLR, and LRRL), which were employed equally often in random fashion on successive trials. At each choice-point, therefore, the animal was required to go right on half the trials and left on half the trials, depending on which of the two lamps was lighted.

In the successive problem, the lamps at each choice-point were either both on or both off. Half the animals of Group II were required to go right to double-light and left to double-dark, while the other half of the animals were required to respond in the reverse manner. As in the simultaneous problem, of course, an incorrect turn led to a barrier, while a correct turn led to subsequent units and finally to the food-box. Again, the stimuli were so arranged that the animals were required

to make two right turns and two left turns on each trial. The same turning orders were employed as for the simultaneous problem, and trials were matched so that animals of both groups were required to make the same sequence of turns on any given trial.

Throughout the experiment the animals were kept on a 24-hour feeding schedule. During the first three days the animals were placed in the food-box and allowed to eat for several minutes. The apparatus was then set up in such a way that only one discrimination-unit separated the starting-box from the food-box, and on the fourth day the groups were given six trials in the one-unit apparatus. The choice-point presented the same arrangements of stimuli to each animal as it would encounter in the main part of the experiment, and responses were reinforced in a manner appropriate to the problem subsequently to be learned by the animal. On the fifth day, training in the four-unit apparatus was begun. Four trials per day were given and each animal was run to a criterion of four successive errorless trials (16/16 correct choices). At the conclusion of each trial the animal was allowed a 20-second period of feeding in the goal-box.

RESULTS AND DISCUSSION

The course of learning in the two groups is shown by the curves of Figs. 2 and 3. As each animal reached the criterion, its training was terminated, but for the purpose of plotting these curves it was assumed to have continued to make perfect runs on succeeding days. Figure 2 shows mean errors per trial per day for each of the two groups, retracing errors being considered along

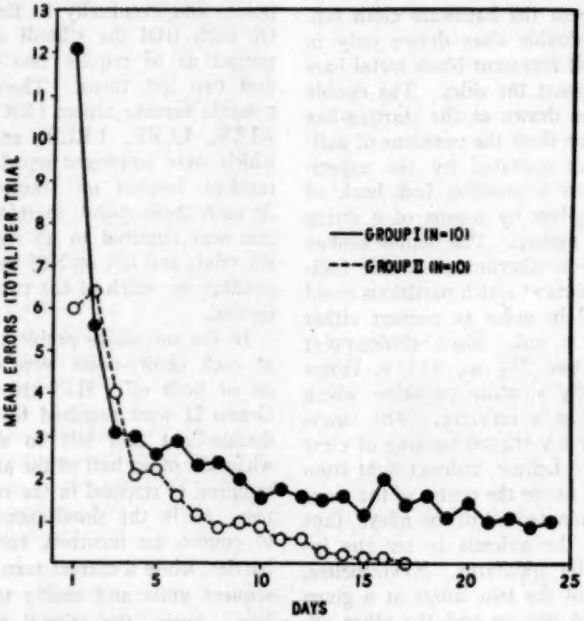

FIG. 2. The course of learning plotted in terms of total errors per trial per day. Group I, simultaneous; Group II, successive.

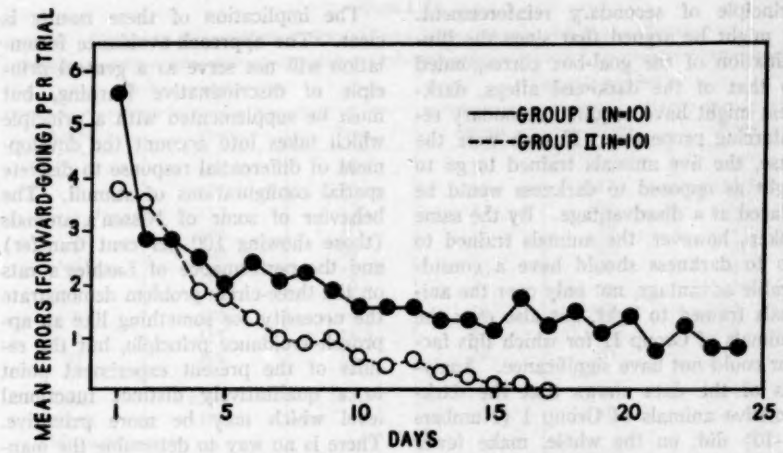

FIG. 3. The course of learning plotted in terms of forward-going errors alone. Group I, simultaneous; Group II, successive.

with forward-going entries into blinds. Both groups retraced considerably during the first few days, which accounts for the fact that the early sections of both curves rise high above two errors per trial, the chance expectancy based on four choice-point encounters. Retracing errors were eliminated early in Group II but continued for some days in Group I. Figure 3 is plotted in terms of forward-going errors alone and shows more clearly the course of learning of the discrimination-problem *per se* in each group. In this curve, again, the higher error score of Group I is in part a function of a greater tendency to retrace, since this group approached the discriminanda more frequently on each trial and thus had greater opportunity for error.

Results for trials and forward-going errors alone, are summarized in Tables 1 and 2. The differences between the two groups are so large that relatively crude, non-parametric statistical techniques are sufficient to indicate their significance. Group II (the successive group) reached the criterion of mastery (one perfect day) in a mean of 48.8 trials and 78.7 errors. The last animal of this group to learn reached the criterion on the 17th day, the same day on which the first animal of Group I reached the criterion. By the 24th day, at which time the experiment was terminated, only three animals of Group I had mastered the simultaneous problem, although all animals in the group were showing signs of learning. The ranking method of Wilcoxon (18) indicates the difference between the two groups in trials and in errors to be significant at better than the one per cent level of confidence.

Even if we consider errors made on the first approach to each choice-point on each trial, the two groups are still sharply separated and the differences in error scores remain significant at better than the one per cent level. From this evidence it seems safe to conclude that the difference in amount of retracing is a *result* rather than a *cause* of the difference in difficulty between the two problems.

A possible explanation of the difference between the simultaneous and successive problems is suggested by the

principle of secondary reinforcement. It might be argued that since the illumination of the goal-box corresponded to that of the darkened alleys, darkness might have acquired secondary reinforcing properties. If such were the case, the five animals trained to go to light as opposed to darkness would be placed at a disadvantage. By the same token, however, the animals trained to go to darkness should have a considerable advantage, not only over the animals trained to light, but also over the animals of Group II for which this factor could not have significance. Analysis of the data shows that the dark-positive animals of Group I (Numbers 6–10) did, on the whole, make fewer forward-going errors than the light-positive animals (Numbers 1–5) of that group, although the difference is not statistically significant. Even the performance of the dark-positive animals, however, was significantly inferior to that of the animals of Group II, at the one per cent level of confidence by Festinger's method (3).

The implication of these results is clear. The approach-avoidance formulation will not serve as a general principle of discriminative learning, but must be supplemented with a principle which takes into account the development of differential response to discrete spatial configurations of stimuli. The behavior of some of Nissen's animals (those showing 100 per cent transfer) and the performance of Lashley's rats on the three-circle problem demonstrate the necessity for something like an approach-avoidance principle, but the results of the present experiment point to a qualitatively distinct functional level which may be more primitive. There is no way to determine the manner in which the simultaneous problem was learned by the animals of Group I which achieved the criterion. If learned configurationally, the greater difficulty of the simultaneous problem may be attributed to the greater similarity of the two stimulus-patterns which it presented to the animal, as the theory of Gulliksen and Wolfle suggests. If mas-

TABLE I

TRIAL SCORES FOR EACH ANIMAL OF EACH GROUP *

Group I (simultaneous)			Group II (successive)		
Animal	Score	Rank	Animal	Score	Rank
1	100	17	11	48	5
2	100	17	12	68	10.5
3	100	17	13	52	6.5
4	96	13	14	36	2
5	100	17	15	44	4
6	100	17	16	52	6.5
7	100	17	17	32	1
8	88	12	18	40	3
9	100	17	19	60	9
10	68	10.5	20	56	8
Total	952	154.5		488	55.5
Mean	95.2			48.8	
Mean difference		46.4**			

* A trial score of 100 was assigned to those animals (Group I) which failed to reach the criterion on the 24th day, at which time training was terminated.

** Significant at better than the one per cent level of confidence (Wilcoxon's method).

TABLE II

FORWARD-GOING ERROR SCORES FOR EACH ANIMAL OF EACH GROUP *

Group I (simultaneous)			Group II (successive)		
Animal	Score	Rank	Animal	Score	Rank
1	164	16	11	69	5
2	157	15	12	117	11.5
3	123	13	13	102	9
4	196	19	14	39	1
5	209	20	15	62	3
6	188	18	16	101	8
7	174	17	17	54	2
8	136	14	18	65	4
9	117	11.5	19	92	7
10	109	10	20	86	6
Total	1573	153.5		787	56.5
Mean	157.3			78.7	
Mean difference		78.6**			

* Error scores are based on 24 days of training irrespective of whether or not an animal reached the criterion on or before this day.

** Significant at better than the one per cent level of confidence (Wilcoxon's method).

tery was based on the acquisition of functional properties by afferent components, this kind of learning may be assumed to involve a more complex, higher order process. In any event, present evidence suggests that two qualitatively distinct processes of learning may operate in discrimination-situations. It is interesting to note the relation between this conclusion and the two-factor conception of modifiability advanced by Maier and Schneirla (11) and by Birch and Bitterman (1). The two processes which seem to operate in discrimination-situations closely resemble the two processes designated by these writers as "sensory integration" and "selective learning." The only discrepancy seems to lie in answers to the question as to which is the more primitive.

The relation between the present results and those of certain previous experiments on the problem of relational learning should also be considered. Lashley and Wade (8) have maintained that opportunity to compare the stimuli to be discriminated facilitates learning. Grice's experiment (4), which was designed to test this hypothesis, failed to show an advantage for the "comparison" group. Another method employed by Saldanha and Bitterman (14) did, however, give a clear advantage to the comparison group under certain conditions. These writers suggest that opportunity for comparison may facilitate learning only when the stimuli to be discriminated are relatively similar. The fact that the simultaneous group of the present experiment, which was provided with opportunity for direct comparison, was surpassed by the successive group, which had no opportunity for direct comparison, may be a function of the great dissimilarity between the two brightness levels employed. It is interesting to ask what would happen in an experiment similar to the present one in which two closely similar brightness levels were employed. It would follow from the position of Saldanha and Bitterman that the simultaneous group should surpass the suc-

cessive group. That is, opportunity for comparison would give an advantage to the former group which would offset the greater fundamental simplicity of the successive problem.

At the outset of the present paper the Gulliksen-Wolfle theory was designated a response-oriented conception in order to differentiate it from the approach-avoidance formulation in the simplest possible manner. Implicit in the discussion of the two orientations, however, was another distinction which may be phrased in afferent terms. For Gulliksen and Wolfle, stimulus-components do not acquire functional significance independently or in isolation, but are integrated to form a functional unity which may comprise spatial as well as other attributes of the part-processes. We might term this intra-situational integrating process *configurational* perception and reserve the term relational perception for the abstractive, trans-situational or trans-contextual process which is emphasized in Lashley's theory. That the two criteria—afferent and efferent—for distinguishing between the formulations with which we began cannot be treated as separate items in a list is suggested by the evidence which Werner has compiled in his stimulating work on mental development (17). Werner, himself, finds it necessary to postulate several genetically distinct discriminative processes, the most primitive of which he refers to as "syncretic" organization in which sensory and motor components are diffusely and inextricably inter-related.

SUMMARY

The question of the relative ease of simultaneous and successive problems has considerable bearing on the validity of current theories of discriminative learning. An experiment is reported in which a group of rats were trained in a multiple discrimination-apparatus to choose the brighter or darker of two alleys (simultaneous problem) while a second group of rats were trained to turn in one direction when both alleys were lighted and in the opposite direction when both were dark (successive problem). The first problem proved to be significantly more difficult than the second. This result is interpreted to mean that under certain conditions the process of discrimination cannot be appropriately described in approach-avoidance terms, but requires the assumption that the animal learns to respond differentially to discrete spatial configurations of stimuli. Taken together, the data presently available suggest the operation of two qualitatively distinct levels of perceptual organization in discriminative learning—a concrete, primitive, intra-situational level and a more abstract, selective, trans-contextual one. The first level is intimately related to motor functions while the second may be largely independent of them.

REFERENCES

1. BIRCH, H. G., & BITTERMAN, M. E. Reinforcement and learning: The process of sensory integration. PSYCHOL. REV., 1949, **56**, 292–308.
2. BITTERMAN, M. E., & COATE, W. B. Some new experiments on the nature of discrimination learning in the rat. *J. comp. physiol. Psychol.*, 1950, **43**, 198–210.
3. FESTINGER, L. The significance of difference between means without reference to the frequency distribution function. *Psychometrika*, 1946, **11**, 97–105.
4. GRICE, G. R. Visual discrimination learning with simultaneous and successive presentation of stimuli. *J. comp. physiol. Psychol.*, 1949, **42**, 365–373.
5. GULLIKSEN, H., & WOLFLE, D. L. A theory of learning and transfer: I. *Psychometrika*, 1938, **3**, 127–149.
6. ——. A theory of learning and transfer: II. *Psychometrika*, 1938, **3**, 225–251.
7. LASHLEY, K. S. An examination of the "continuity theory" as applied to discriminative learning. *J. gen. Psychol.*, 1942, **26**, 241–265.

8. ——, & WADE, M. The Pavlovian theory of generalization. PSYCHOL. REV., 1946, 53, 72–87.
9. LAWRENCE, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *J. exp. Psychol.*, 1949, 39, 770–784.
10. ——. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *J. exp Psychol.*, 1950, 40, 175–188.
11. MAIER, N. R. F., & SCHNEIRLA, T. C. Mechanisms in conditioning. PSYCHOL. REV., 1942, 49, 117–134.
12. MUENZINGER, K. F., & GENTRY, E. Tone discrimination in white rats. *J. comp. Psychol.*, 1931, 12, 195–205.
13. NISSEN, H. W. Description of the learned response in discrimination behavior. PSYCHOL. REV., 1950, 57, 121–131.
14. SALDANHA, E., & BITTERMAN, M. E. Relational learning in the rat. *Amer. J. Psychol.*, 1951, 64, 37–53.
15. SPENCE, K. W. The nature of discrimination learning in animals. PSYCHOL. REV., 1936, 43, 427–449.
16. STONE, C. P. A multiple discrimination box and its use in studying the learning ability of rats: I. Reliability scores. *J. genet. Psychol.*, 1928, 35, 557–573.
17. WERNER, H. *Comparative psychology of mental development*. New York: Harper & Bros., 1940.
18. WILCOXON, F. *Some rapid approximate statistical procedures*. Stamford, Conn.: American Cyanamid Co., 1949.

[MS. received August 11, 1950]

CONDITIONING AND CONDITIONALITY (DISCRIMINATION)

BY O. H. MOWRER

University of Illinois

AND

R. R. LAMOREAUX

State Hospital, No. 3, Nevada, Missouri [1]

Anyone who has attempted to study learning by the conditioned-response technique has encountered the complication of having responses occur, not only when the so-called conditioned stimulus is presented, but also in the intervals between trials. Such "spontaneous," or interval, responses are objectionable in that they make it less easy to determine the precise extent of conditioning; for if a response is likely to occur in the absence of the CS, the occurrence of the response when the CS is present involves no assurance that the CS has necessarily *caused* the response—it might have occurred anyway.

Various methods have been proposed for minimizing the occurrence of these interval responses; and, once such precautions as seem feasible have been taken, the tendency is henceforth to pay no more attention to them. However, in a number of prior studies (3, 10, 13, 14) the present writers have systematically recorded the incidence of interval responses throughout the course of conditioning, and the resulting data suggest that many so-called conditioning studies (including a number of our own) have not dealt with conditioning in pure form but rather in a form which is a complex admixture of conditioning and *discrimination learning*. If this is true, it follows that these studies give a misleading picture of conditioning, since discrimination is a relatively slow process which, if confounded with conditioning, would have the effect of materially depressing the resulting measures.

More concretely, our supposition is that in the usual type of conditioning experiment true conditioning occurs very quickly but that the measures used are such as to call for a considerable amount of added discrimination learning. Let us assume, for example, that we have designed an experiment with rats as subjects in which the conditioned stimulus is a buzzer, the unconditioned stimulus an electric shock administered from a floor-grill, and the recorded response a leap into the air. After one or, at most two, paired presentations of the buzzer and shock, a marked change will be noted in an animal's behavior and appearance; generalized activity will be much reduced, ears will be laid back, eyes will protrude, and breathing will be labored. In common parlance the rat is *afraid;* and if there is any doubt that genuine learning has occurred, it can be dispelled by removing the animal from the apparatus for several days, or even weeks, and then returning it. It will be at once apparent that the rat "remembers" its earlier experience in the apparatus, for it will now behave very unlike a naïve rat

[1] The research on which this paper is based was carried out while the writers were at Harvard University. Grateful acknowledgment is here made for facilities and financial support made available by the Laboratory of Social Relations.

which is put into the apparatus for the first time. The rat will again show signs of fear; and we must conclude that the fear reaction, as a result of only one or two occurrences of the shock in the presence of experimental-situation-plus-buzzer, has become connected, not only to *that* stimulus-compound, but also (through generalization) to experimental-situation-without-buzzer. The *fear reaction*, now elicited by the situation-as-a-whole (whether with or without the specific "conditioned stimulus") is, we suggest, the *true conditioned response*. Conditioning, so conceived, is a type of learning that occurs *very* rapidly, often in as few as one or two trials.

But in the interest of objectivity we have agreed that we shall use as an indicator of conditioning, not the diffuse changes noted above, but the more striking, explicit response of leaping into the air. Although there is an initial, relatively low probability that a fearful, confined rat will leap into the air, this probability is not high enough, in the beginning, for this response to serve as a very reliable index of fear. However, the connection between fear and leaping can be rapidly strengthened by following a so-called "instrumental" conditioning procedure, which would allow the leaping, when it eventually occurs, to carry the rat completely out of the experimental situation. Since it is the situation-as-a-whole which has become the conditioned stimulus for the fear reaction, any response which will remove the rat from that situation will be powerfully reinforced by fear-reduction. In this way the leaping can be differentially strengthened and made to rise rapidly in the rat's original hierarchy of responses to the acquired drive of fear.

It will be apparent that in using an overt skeletal response such as leaping as an index of fear, we have introduced into our study of conditioning the second fundamental type of learning, namely, problem-solving. Elsewhere (11, 12) an extended argument has been advanced for believing that there are two basically different forms of learning: conditioning, which occurs on the basis of a pure contiguity principle, and problem-solving, which involves reinforcement in the sense of drive-reduction, reward, satisfaction. Although, in the hypothetical situation just described, we somewhat complicate our measurement of conditioning by interposing a certain amount of problem-solving learning, the net effect is not too serious, since the requisite problem-solving learning is simple, occurs rapidly, and does not unduly becloud the picture which we get of the antecedent process of conditioning. Anyone who has allowed a rat to use leaping as a means of getting out of an experimental situation in which electric shock has been administered knows that this response will become "stamped in" very rapidly. Therefore the use of this response as an indicator of the process whereby the intermediate conditioned response-drive of fear is acquired introduces no very serious distortions.[2]

Let us now consider the procedure more typically employed in laboratory investigations of conditioning. If, un-

[2] Cattell (2), McDougall (8), Woodworth (19), and others have criticized the common stimulus-response, or $S \rightarrow R$, conception of behavior for omitting the *organism* and have substituted the formula: $S \rightarrow (O) \rightarrow R$. The above analysis fills in the details further by noting that the S (buzzer) produces, *within* the organism, the response (R_f), which gives rise to the experience, or stimulus, of fear (S_f), which in turn produces the observed external jumping behavior (R_j). The formula would thus read: $S_b \rightarrow (R_f \rightarrow S_f) \rightarrow R_j$. The two-factor conception of learning alluded to above posits that the $S_b \rightarrow R_f$ connection is acquired through *conditioning* and the $S_f \rightarrow R_j$ connection is based on *problem-solving*. To speak simply of the $S_b \rightarrow R_j$ "connection" is to omit, as Tolman has aptly observed, a number of "intervening variables" (18).

der circumstances already described, one has decided to use the response of leaping as an indicator of fear-learning, the usual policy is this. The experimenter will not allow the leaping response to remove the rat physically from the experimental situation; but if the leaping occurs when the buzzer is on, the experimenter may terminate the buzzer.[3] In a manner of speaking, the rat thus "escapes" from situation-with-buzzer into situation-without-buzzer. And since, for reasons previously noted, the reaction of fear will be more strongly conditioned to situation-with-buzzer than to situation-without-buzzer, termination of the buzzer will have the effect of reducing the rat's fear at least a little and will thus reinforce, at least mildly, the leaping behavior.

But fear will continue to be present between trials; and when the rat tries likewise to deal with this fear by leaping, nothing happens—the fear is not reduced. There will thus be a tendency for the response of leaping-to-fear between trials to extinguish and for this extinction to generalize to leaping-to-fear in situation-plus-buzzer. However, when leaping fails to occur in response to situation-plus-buzzer, the shock follows; and the connection between situation-plus-buzzer and fear is reinforced. The intensified fear of situation-plus-buzzer will revive the leaping response, which will be reinforced by buzzer-termination (fear-reduction); and this response will tend, once again, to generalize to fear of situation-without-buzzer. By this process of generalization, extinction, and differential reinforcement, situation-with-buzzer gradually becomes identified as *"dangerous"* and situation-without-buzzer as *"safe."*[4]

When this point has been reached, our subjects will regularly leap into the air when the buzzer is presented and only, or at least predominantly, at this time. But how misleading it is to say: "Conditioning has now occurred!" Conditioning, in the most primitive and fundamental sense of the term, occurred much earlier, and what has now been accomplished is a relatively refined and advanced stage of discrimination learning.[5]

[3] We have elsewhere (13) made a systematic comparison of the results obtained by using a conditioned stimulus of fixed duration vs. one of variable duration. Our findings indicate that, at least under the specified experimental conditions, the variable CS was decidedly superior. See also the results of an experiment by Traum, Horton, and Mowrer (12).

[4] As indicated in footnote 1, there is a two-link connection between situation-plus-buzzer and the response of jumping, the first of which is $S_b \rightarrow R_f$ and the second, $S_f \rightarrow R_j$. These two connections are assumed to represent two forms of learning, conditioning and problem-solving, respectively. The question now arises: at which level does the discrimination take place, at the conditioning or the problem-solving level. In other words, we want to know whether discrimination, in an experiment of this kind, consists of the rat's finding out merely that jumping in response to fear does no good in situation-without-buzzer and is effective only when buzzer is present, or does the discrimination consist of a rat's "finding out" that situation-without-buzzer really is not dangerous and that it is only when buzzer is present that fear is justified? Or perhaps some degree of discrimination occurs at *both* levels!

[5] When one views the end product of this type of training, it is perhaps natural to refer to buzzer as *the* conditioned stimulus, since it seems to be the sole and sufficient cause of the leaping behavior. But this impression can be quickly dissipated by taking a rat so trained out of the experimental apparatus, putting it on the floor, and then sounding the buzzer: The likelihood that the buzzer will now elicit a leap is very slight, thus indicating that it is really buzzer-plus-experimental-situation which is the conditioned stimulus. By associating the buzzer with shock in a great many different situations, one can ultimately reach the point where leaping will occur to buzzer-plus-any-situation, which is to say to buzzer *alone*. The buzzer can now be termed *the* stimulus for leaping (or, more accurately, for the fear that produces leaping); but this state of affairs is the end product of extended training,

Although the foregoing analysis is by no means complete, we have perhaps proceeded far enough for it to be evident that, in the example chosen and in many comparable experiments, we are dealing with an extrardinarily complicated picture, involving not only the two basic processes of conditioning and problem-solving but also the more complex phenomenon of discrimination learning. Indeed, we may conjecture that the learning curves obtained in many conventional studies of conditioning are mainly discrimination curves. If this be true, we can then predict that such curves will be prominently influenced by varying the difficulty with which the neutral interval between trials (situation-without-CS) can be differentiated from the "live," meaningful period when the CS is in effect. In other words, by varying factors which we would expect to affect discrimination learning, we can perhaps obtain more direct evidence of the extent to which, in many so-called conditioning experiments, we are really dealing with discrimination.

With this reasoning in mind, the following experiment has been performed.

not an early development. That conditioned responses tend to become connected to the experimental-situation-as-a-whole and that change in that stimulus-totality tends to reduce the intensity or likelihood of the conditioned response is indicated by results recently reported by Lichtenstein. In concluding his observations in this connection, Lichtenstein says: "Feeding inhibitions [based on a conditioned fear reaction] may be temporarily relieved in some animals by altering one or more aspects of the experimental situation. Some animals will eat when the experimenter is present in the room [which he is not during experimentation—cf. similar observations by Pavlov (16)] and some will eat when the harness is removed. Other dogs, however, cannot be made to eat in the experimental room even under considerably altered conditions" (7, p. 29).

SUBJECTS AND APPARATUS

The subjects used in this experiment were 24 Lashley-strain male rats approximately four months of age. They were randomly divided into two equal groups and submitted, respectively, to the two procedures described below.

The apparatus here employed has been fully described by Mowrer and Miller (15). It consists essentially of a rectangular box, with a glass front and a metal-grill floor which can be electrically energized.

For one group of subjects the CS, so-called, consisted of the sound normally produced by an electric bell. For the other group the CS was produced by cutting out the interrupter circuit of the bell, which meant that when the experimenter pressed the key controlling the bell, a single, gong-like sound was produced. A device for damping the reverberation of the bell insured that the CS was, in the second case, virtually instantaneous in duration. For both groups of animals, the UnCS was produced by a 60-cycle alternating current, 140 volts, with a limiting series resistance of 250,000 ohms.

PROCEDURE

In this experiment the overt response investigated was a vertical jump. The criterion for this response was merely that the animal got all four feet off the grill simultaneously, but the response ordinarily consisted of a leap which carried the animal several inches into the air.

In the case of both groups of subjects, an "instrumental" rather than "classical" conditioning procedure was followed. The CS was presented and if, after five seconds, there was no leaping response, it was followed by the UnCS; but if a leap occurred with the five-second interval, the UnCS was omitted. On those trials when the UnCS had to be applied, it remained

on, not for a fixed period of time, but until the leaping response had been elicited. We thus employed what might be called a double-instrumental procedure: the leaping response, if it occurred in advance of the shock, served to avert the shock; and if the shock was not thus averted, the leaping response could terminate it (13).

The only difference between the procedures employed with the two groups was thus in the nature of the CS. In the case of the Group-I animals, where the CS was continuous rather than instantaneous, the occurrence of leaping was the occasion for terminating the CS. If, in this group, the leaping did not occur, the CS continued until the UnCS came on and had elicited the required reaction; the CS and the UnCS then terminate together. Since, in the Group-II procedure, the CS was instantaneous, there was no question as to when it should be terminated. If, however, jumping occurred five seconds after the instantaneous CS occurred, the UnCS was omitted, just as in the Group-I procedure.

All animals received ten training trials per day, at regular one-minute intervals, and all were run for ten consecutive days. "Spontaneous" jumping responses were recorded between trials and for one minute before the first trial and for one minute after the last trial.

It should be noted that in the Group-I procedure, the situation-when-dangerous (bell present) was easily distinguished from situation-when-safe (bell absent). Or, in Gestalt terms, "figure" stood out clearly against "ground." But in the Group-II procedure, the differentiation was more difficult. The situation-when-safe was always quiet (as it was in the Group-I procedure); but situation-when-quiet was *not* always safe, since silence occurred during both the inter-trial (safe) interval and also during the inter-stimulus (dangerous) interval. The silent interval between trials could thus be discriminated, in the Group-II procedure, from the silent intervals between stimuli solely on the basis of the *stimulus trace* ("immediate memory") produced by the momentary, gong-like CS.[6]

RESULTS AND DISCUSSION

Figure 1 shows graphically the average number of leaping responses made during the five-second interval following onset of the CS by the two groups of animals on each of the ten days of experimentation. These responses are what would ordinarily be termed "conditioned responses." The average number of such responses made by the Group-II animals during the whole experiment, on all (100) trials, was only 49.7, whereas for the Group-I animals it was 78.9.

These results may seem not at all surprising; they may appear merely to confirm the finding, reported by various other experimenters, that "trace" conditioning is more difficult than is simultaneous conditioning. Trace conditioning, first defined by Pavlov (16), refers to the acquisition of a CR in a situation wherein the CS comes on and then goes off a considerable time before the UnCS occurs. In simultaneous conditioning the CS comes on and remains on until the UnCS occurs (or perhaps until it terminates). According to these definitions, our Group-II procedure involved trace conditioning, and the Group-I pro-

[6] A similar, though less marked, difference in the procedures employed with the two groups might have been provided by having an auditory CS in one case loud (or a visual CS bright) and in the other case weak (dim). As is well known, an intense CS produces better "conditioning" than does a less intense one. There may be a direct, dynomogenic effect here; but there is also the possibility that the more intense the CS the more readily situation-with-CS is differentiated from situation-without-CS.

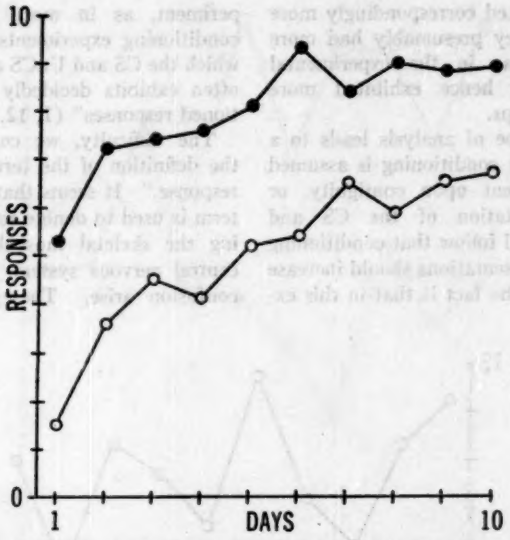

FIG. 1. Average number of leaping responses elicited by a conditioned stimulus in the 12 animals comprising Group I (filled circles) and in the 12 animals comprising Group II (open circles). The only difference in the procedure followed with the two groups of subjects was with respect to the ease with which the conditioned stimulus could be differentiated from the interval between trials. Since this variable produced a conspicuous difference in the amount of "conditioning" obtained, the inference is confirmed that would-be measurement of conditioning is commonly contaminated by measurement of discrimination learning.

cedure involved simultaneous conditioning. Therefore, the difference in results may, without further analysis, be interpreted as simply demonstrating a phenomenon which is already well established.

But interpretation of our results in the manner just suggested does not readily accommodate the findings shown in Fig. 2. Here it will be noted that the incidence of leaping responses which occurred in the interval between trials was, not *less*, but *much greater* in the Group-II animals than in the Group-I animals. We see nothing in traditional conditioning theory which would either predict or explain this finding, that a trace-conditioning procedure gives rise to more inter-trial responses than does the simultaneous-conditioning procedure.

One fact is immediately apparent: in an instrumental conditioning procedure of the kind here employed, there will always be a reciprocal relationship between the number of so-called conditioned responses exhibited by the subjects and the number of times the UnCS impinges upon them. That is to say, in the present study whenever a rat jumped within the five-second interval following onset of the CS, the UnCS (shock) was prevented from occurring. This fact means that since the Group-I animals jumped relatively often in response to the CS they were shocked fewer times. They presumably had as a result less generalized fear in the experimental situation and hence exhibited fewer inter-trial jumps. By contrast, the Group-II animals, which jumped less often in response to the

CS, were shocked correspondingly more frequently; they presumably had more generalized fear in the experimental situation, and hence exhibited more inter-trial jumps.

But this type of analysis leads to a paradox: Since conditioning is assumed to be dependent upon contiguity, or paired presentation of the CS and UnCS, it would follow that conditioning and paired presentations should increase *pari passu*. The fact is that in this experiment, as in most "instrumental" conditioning experiments, the group in which the CS and UnCS are paired most often exhibits decidedly fewer "conditioned responses" (1, 12, 13).

The difficulty, we conjecture, is in the definition of the term "conditioned response." It seems that whenever this term is used to denote responses involving the skeletal musculature and the central nervous system, ambiguity and confusion arise. These responses are

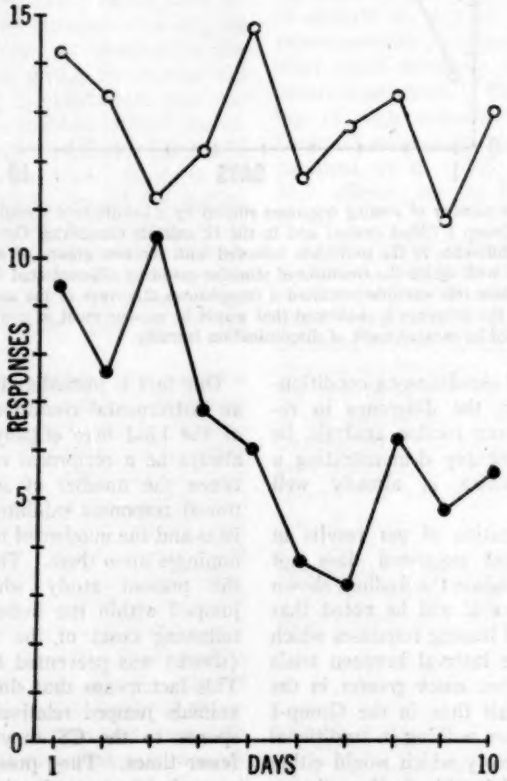

FIG. 2. Average number of spontaneous, or interval, responses made by the animals in Group I (filled circles) and by the animals in Group II (open circles). It will be noted that there is thus a reciprocal relationship, in this type of experiment, between the number of responses made to the conditioned stimulus (see Fig. 1) and the number of similar responses made in the interval between trials. This finding presents a paradox for certain forms of conditioning theory.

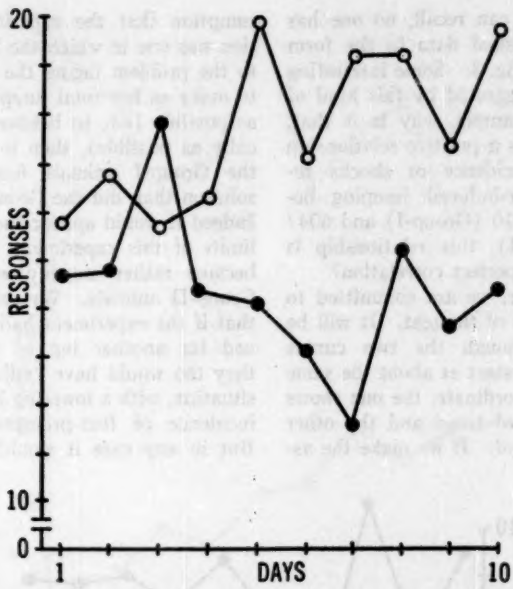

FIG. 3. Average number of leaping responses made by the Group-I animals (filled circles) to both the conditioned stimulus and in the intervals between trials and the average number of such responses made by the Group-II animals (open circles). Since the number of responses which any given animal needed to make per day in order to avoid receiving the unconditioned stimulus (shock) was only 10, it is clear that the declining curve for the Group-I animals means that these animals were making a better over-all "adjustment" to the experimental situation than were the Group-II animals, with their ascending curve. If the experiment had lasted long enough, the Group-II animals would, we believe, have made an almost equally good adjustment, but within the limits of the experiment the difference was striking.

better referred to as problem-solving responses, with the term "conditioned response" being reserved for responses involving smooth muscle and glands and the autonomic nervous system (11, 12).

With this revised terminology in mind, the paradox mentioned above can be resolved in the following manner. In Fig. 3 we have reproduced two composite curves. One of these (the lower) was derived by adding the number of jumps the Group-I animals made in response to situation-with-buzzer and the number of jumps they made in response to situation-without-buzzer; and by following the same procedure with respect to the Group-II data, we obtained the other curve here shown. Now if we make the not unreasonable assumption that there is a roughly one-to-one relationship between the total amount of fear experienced by our subjects and the total amount of jumping they did during the course of the experiment, the paradox disappears: since the Group-II animals received the greater number of shocks (604), we are not surprised to discover that they showed the greater over-all incidence of fear-produced jumping (2,121); and since the Group-I animals received the smaller number of shock (253), we are prepared to find that they showed the smaller over-all incidence of jumping (1,710).

So far as we can recall, no one has previously published data in the form represented in Fig. 3. Some interesting problems are suggested by this kind of analysis: for example, why is it that, although there is a positive relationship between the incidence of shocks received and fear-induced jumping behavior—253/1,710 (Group-I) and 604/2,121 (Group-II), this relationship is by no means a perfect correlation?

Here, however, we are committed to our original line of thought. It will be noted that although the two curves shown in Fig. 3 start at about the same position on the ordinate, the one shows a definite upward trend and the other a downward trend. If we make the assumption that the experimental situation was one in which the ideal solution to the problem facing the subjects was to make as few total jumping responses as possible (*i.e.*, to behave as economically as possible), then it is clear that the Group-I animals found a better solution than did the Group-II animals. Indeed it would appear that, within the limits of this experiment, the situation became rather steadily worse for the Group-II animals. We may conjecture that if the experiment had been continued for another ten or twenty days, they too would have "adjusted" to the situation, with a lowering in the over-all incidence of fear-prompted responses. But in any case it would appear that

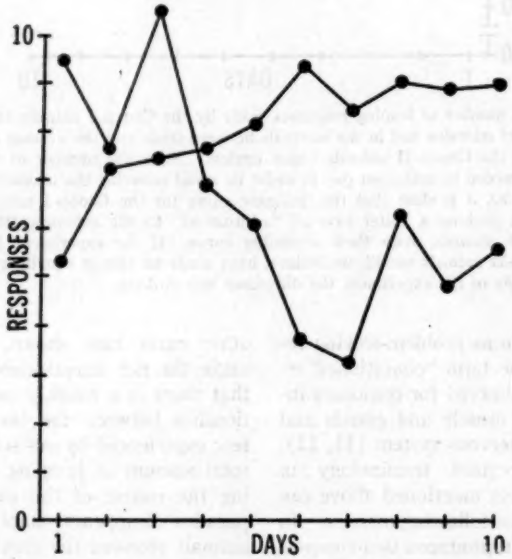

FIG. 4. Average number of responses made to the conditioned stimulus (ascending curve) and the average number of responses made between trials (descending curve) by the animals in Group I. The purpose of bringing these two curves into juxtaposition on the same graph (they have already appeared in Figs. 1 and 2) is to high-light the extent to which discrimination, rather than conditioning only, occurred in this experiment. If experimental-situation-with-CS is thought of as "positive" (in the sense of a conventional discrimination experiment) and if experimental-situation-without-CS is thought of as "negative," then it will be seen that the increment in response to the former and the decrement in response to the latter fulfill the usual pattern of expectation with respect to discrimination learning.

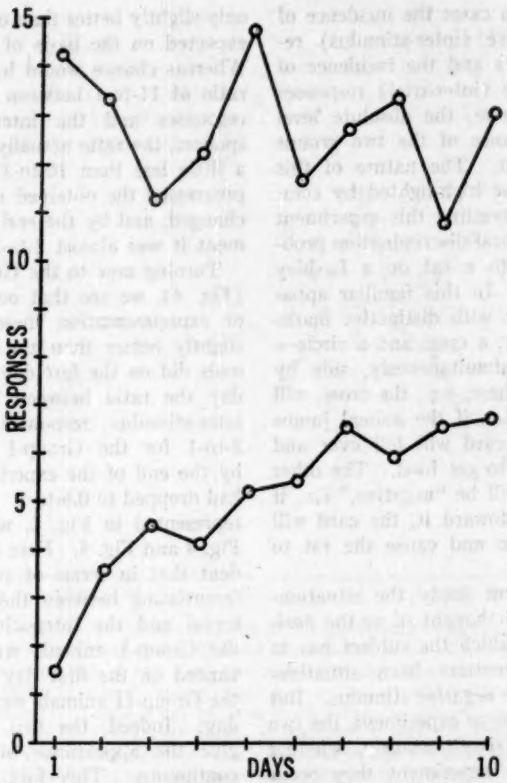

FIG. 5. Average number of responses made to the conditioned stimulus (ascending curve) and the average number of responses made between trials (descending curve) by the animals in Group II. Although the responses to the "positive" stimulus (situation-with-CS) increases and the responses to the "negative" stimulus (situation-without-CS) decreases (thus showing the same pattern or trend as represented in Fig. 4), the absolute levels of the two curves are quite different. This is presumably a reflection of the difference in the level of difficulty of the discrimination task which faced the Group-I animals and the Group-II animals.

the Group-II animals were being called upon to make a much more difficult kind of adjustment than were the Group-I animals and that the essence of this difference was the relative ease with which the inter-stimulus and the inter-trial intervals could be discriminated in the two cases.

That we are here dealing with a problem of discrimination is further suggested by Fig. 4 and Fig. 5. The curves which appear in these two graphs are the same as those which appear in Figs. 1 and 2, but now the inter-stimulus (so-called "conditioned") jumping responses and the inter-trial ("spontaneous") responses for Group-I appear together (Fig. 4) and the inter-stimulus and inter-trial responses for Group-II appear together (Fig. 5).

In both Fig. 4 and Fig. 5 there is clear indication of discrimination learn-

ing; *i.e.*, in both cases the incidence of relevant, effective (inter-stimulus) responses increases and the incidence of useless, abortive (inter-trial) responses declines. However, the absolute level of the performance of the two groups is very different. The nature of this difference can be high-lighted by comparing and contrasting this experiment with a hypothetical discrimination problem presented to a rat on a Lashley jumping-stand. In this familiar apparatus two cards with distinctive markings—let us say, a cross and a circle—are presented simultaneously, side by side. One of these, *e.g.*, the cross, will be "positive," *i.e.*, if the animal jumps toward it, the card will fall over and permit the rat to get food. The other card (circle) will be "negative," *i.e.*, if the rat jumps toward it, the card will remain in place and cause the rat to get a mild fall.

In the present study the situation-with-CS can be thought of as the *positive* stimulus, which the subject has to learn to differentiate from situation-without-CS, the *negative* stimulus. But in the Lashley-type experiment, the two stimuli occur *simultaneously*, whereas in the present experiment they occur *successively*. And this fact leads to another important distinction. In the beginning of a Lashley-type experiment, as the right-left position of the two cards varies randomly from trial to trial, the chance that a rat will jump to one card or the other is 50–50. But in the present experiment, on the basis of chance alone, a rat was eleven times more likely, in the beginning, to respond to the "negative" stimulus than to the positive one, since the inter-stimulus interval (situation-without-CS) lasted 55 seconds and the inter-stimulus interval (situation-with-CS) lasted for only 5 seconds.

On the first day of training the Group-II animals (Fig. 5) responded only slightly better than one might have expected on the basis of chance alone. Whereas chance would have dictated a ratio of 11-to-1 between the inter-trial responses and the inter-stimulus responses, the ratio actually obtained was a little less than 10-to-1. As training progressed the obtained ratio gradually changed, and by the end of the experiment it was almost 2-to-1.

Turning now to the Group-I animals (Fig. 4), we see that on the *first* day of experimentation these animals did slightly better than the Group-II animals did on the *last* day. On the first day the ratio between inter-trial and inter-stimulus responses was almost 2-to-1 for the Group-I animals; and by the end of the experiment the ratio had dropped to 0.6-to-1. These facts are represented in Fig. 6, which combines Fig. 4 and Fig. 5. Here it becomes evident that in terms of successfully differentiating between the inter-trial interval and the inter-stimulus interval, the Group-I animals were as far advanced on the first day of training as the Group-II animals were by the tenth day. Indeed, the two sets of curves give the appearance of being almost continuous. This fact is largely coincidence, but it serves to emphasize the difference in the *level of difficulty* of the discrimination problems which the two groups faced.

This type of analysis shows that although the discrimination problem to which the Group-I animals were subjected was relatively easy, it was nevertheless a discrimination problem—and did not provide a clear, uncomplicated measure of *conditioning*. In fact, if we look back at Fig. 3, we may surmise that in Group I the conditioning, properly speaking, was at its height on the *first day* of the experiment, for there were more fear-produced jumps on this day than any time thereafter. This observation tends to confirm our thesis

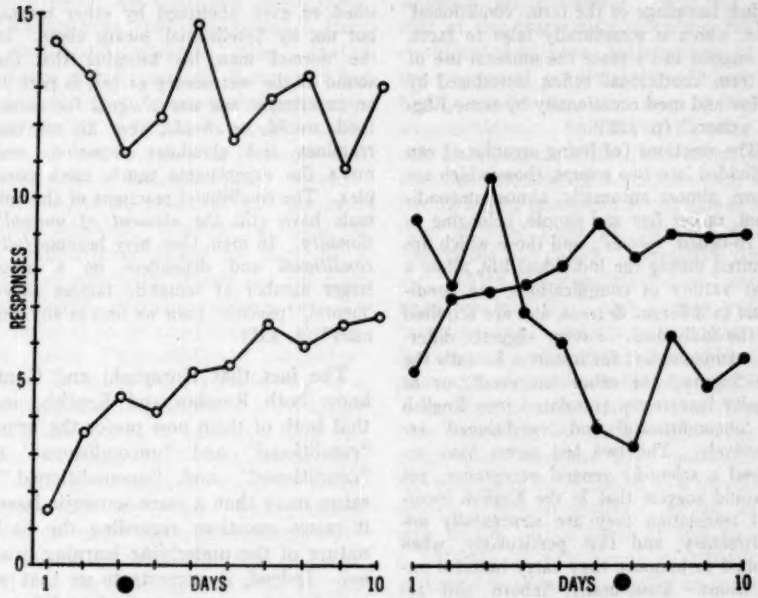

FIG. 6. Average number of responses made by the Group-I animals (at the right) and by the Group-II animals (at the left) to the conditioned stimulus and in the inter-trial intervals. This figure is a composite of Figs. 4 and 5, so arranged as to show that the performance of the Group-I animals began at a level roughly equivalent to that at which the Group-II animals ended their performance. This is a way of again emphasizing the difference in the difficulty of the learning task which the two groups faced. Incidentally, it may noted that the curves for the responses to the conditioned stimulus are much smoother than are those for the interval responses. Why this should have been the case is not immediately evident.

that conditioning, properly defined, is a form of learning which may occur very quickly, and that many conventional methods of studying this phenomenon mask and overlay it by imposing special circumstances which are irrelevant to the study of conditioning in the strict sense of the term.

TERMS AND DEFINITIONS

When this paper was originally written, it was given the title: "Further Analysis of the Concept of 'Conditioning.'" Later a paper by Gantt (4) came to our attention which suggested the title actually used and prompted the line of thought which will be developed in this concluding section.

In the paper cited Gantt uses the terms "conditional" and "unconditional" in place of the commoner terms, "conditioned" and "unconditioned" to characterize the stimuli and the responses connected with conditioned-response learning. This usage recalled to us a not much cited but suggestive chapter in Korzybski's book, *General Semantics* (6), in which the author raises, by implication, the question of the relationship between *conditioning* and *conditionality* (which, as the following excerpts will show, is essentially a synonym for discrimination). Korzybski remarks:

"Before going further, I will analyze and suggest the complete elimination from the

English Language of the term 'conditioned' reflex, which is structurally false to facts, and suggest in its place the uniform use of the term 'conditional' reflex, introduced by Pavlov and used occasionally by some English writers" (p. 328).

"The reactions [of living organisms] can be divided into two groups, those which are inborn, almost automatic, almost unconditional, rather few and simple, belonging to the so-called 'species'; and those which are acquired during the individual life, allow a great variety of complications, are *conditional* in different degrees, and are acquired by the individual. Pavlov suggests different terminologies; for instance, he calls the one 'inborn,' the other 'acquired'; or as usually incorrectly translated into English as 'unconditioned' and 'conditioned' respectively. The two last terms have received a scientific general acceptance, yet I would suggest that in the English incorrect translation they are *structurally unsatisfactory*, and that particularly, when applied to humans, they carry harmful implications. Structurally, 'inborn' and 'acquired' are entirely satisfactory. Terms like 'conditional' and 'unconditional' (in the original language of Pavlov), although less satisfactory, are more appropriate [than 'conditioned' and 'unconditioned'], as they do not imply some sort of 'causelessness.' In fact, the 'unconditioned' salivary reactions *are conditioned* and produced by the physiochemical effect of the food, and so to call them 'unconditioned' is structurally erroneous. The terms 'conditional' and 'unconditional' do not have similar implications, and carry others, as for instance, the possibility of very important *degrees of conditionality*, establishing the ∞-valued character of the reactions; conditional meaning non-absolute, and non-one-valued.

"For these structural reasons, I shall use the term 'inborn' and 'acquired' or else 'unconditional' and 'conditional' reactions" (pp. 332–333).

"Once a conditional reaction is established with an animal, no amount of any sort of 'intellectual' persuasion, or the like, would disturb his glandular secretions, as the animal's range of 'meanings' is very limited. These secretions can be diminished or even abolished by other means, but not by 'intellectual' means alone. In the 'normal' man, his 'knowing' that the sound of the metronome or bell is part of an experiment and not a signal for actual food, would, *or should*, alter his nervous reactions and glandular secretions and make the experiments much more complex. The conditional reactions of the animals have still the *element of unconditionality*. In man, they may become *fully conditional* and dependent on a much larger number of semantic factors called 'mental,' 'psychic,' than we find in any animals" (p. 335).

The fact that Korzybski and Gantt know both Russian and English, and that both of them now prefer the terms "conditional" and "unconditional" to "conditioned" and "unconditioned,"[7] raises more than a mere semantic issue: it raises questions regarding the basic nature of the underlying learning process. Indeed, it suggests to us that we may have here, not merely two different sets of terms for referring to the same phenomenon, but rather *two phenomena* and that possibly the confusion about terms reflects some measure of ambiguity with respect to the phenomena themselves.

As an expository device let us use the letters, CS and UnCS, to stand, not for any particular words, but for the actual stimulus events involved in a conditioning experiment; and let us use CR and UnCR to refer, in like manner, to the responses elicited, respectively, by the two types of stimulus events. The UnCR, then, is commonly a response which is literally *unconditional, i.e.*, it occurs when the UnCS is presented, without regard to context, time or place; that is, it generalizes widely and is not discriminant. The term "unconditioned" does not carry these connota-

[7] In his 1928 translation of Pavlov's *Lectures on Conditioned Reflexes* (17) Gantt consistently employed the terms "conditioned" and "unconditioned."

tions, which are valid and useful ones; it merely says, "not conditioned." Here a number of ambiguities arise. In the passage already quoted, Korzybski has called attention to one unfortunate connotation of the term "unconditioned"; certainly we do not wish to imply that an UnCR is indeterminate, causeless. Likewise, if we speak of an "unconditioned stimulus," it might be taken, on a commonsense basis, to refer to a stimulus which has simply not yet been conditioned, implying perhaps that the stimulus is *to be* conditioned. Finally, the term "unconditioned stimulus" might be interpreted as referring to a stimulus that has been but is no longer conditioned, *i.e.*, a stimulus that has been "extinguished" (deconditioned). All things considered, then, it would appear that the terms "unconditioned stimulus" and "unconditioned response" are indeed "structurally false to facts" and can be usefully replaced by the terms which Korzybski and Gantt propose.

When we come to the question of what terms we may most meaningfully apply to the CS and the CR, the case is more complicated but also more instructive. Perhaps the best way to cut into the heart of the issue is to note that a *conditioned* response may have various degrees of *conditionality*. As already indicated in the main body of this paper, a CR (such as fear) may be highly precise, highly conditional; for example, it may occur in the experimental setting when and only when the specific CS is present. Said otherwise, the CR will occur to experimental-situation-with-CS but not to experimental-situation-without-CS and not to CS-without-experimental-situation. On the other hand, a CR may be relatively *unconditional*; that is to say, if the CS has been reinforced in a wide variety of situations, the response tends to generalize widely and to occur in all, or virtually all, situations. We thus arrive at the interesting inference that a *conditioned response* may possess varying degrees of *conditionality*: it may be highly *conditional*, or it may be highly *unconditional*.[8] And if this be true, we can hardly escape the final conclusion that conditioning and conditionality are two different phenomena, the one referring to that form of learning which Thorndike has called "associative shifting," which Hollingworth called "redintegrative learning," and Pavlov called "conditioning," and the other term referring to the related but distinguishable phenomenon of discrimination learning.

Perhaps we can sharpen this discussion somewhat by noting that when a conditioned response first develops, it has what may be regarded as a middling degree of conditionality. As we have seen in the earlier parts of this paper, a CR will at first occur to situation-without-CS almost as readily as to situation-with-CS. However, the CR is not likely to occur in widely different situations, regardless of whether they do or do not contain the specific CS. But by means of further training, which in the one case is "discrimination training" and in the other case is "generalization training," we can then push the CR in either of the two directions along the conditionality continuum. This reasoning suggests, then, the possibility that we may wish to speak, on occasion, of highly conditional conditioned responses and, perhaps, of highly unconditional ones!

This analysis leads to a number of rather intriguing deductions, but only one of them will be cited here. Pavlov

[8] This distinction is implicit in Korzybski's remark, quoted above, that "the conditional [conditioned] reactions of the animals have still the *element of unconditionality*." Obviously this author is here having to use the same term to refer to two different phenomena. Our proposal with respect to terminology avoids this difficulty.

and others have shown that, in animals, "higher-order" conditioning cannot ordinarily be carried beyond the third or fourth order. We may now interpret this finding as meaning that as we go from the original UnCS to the first-order CS, there is considerable loss of unconditionality so that the latter is only partially effective as an "UnCS" in the second-order conditioning. And with each higher level of conditioning, or would-be conditioning, there is further loss of "unconditionality," until at length conditioning fails to occur at all. Basic theory holds that conditioning is dependent upon the paired presentation of a CS and an UnCS, where the UnCS possesses a considerable degree of unconditionality (if for no other reason than the fact that the results produced by the UnCS must not be significantly changed by the addition of the CS in the conditioning procedure). If we try, then, to obtain conditioning by combining a CS and a so-called UnCS which possess only negligible unconditionality, it is not surprising if conditioning fails to occur. (Perhaps if a higher-order or derived UnCS has but a minimum of unconditionality, *i.e.*, little capacity to withstand generalization of circumstances, then the mere addition of the new CS will suffice to "change the situation" sufficiently so that the UnCS will no longer produce the relevant response.)

On the other hand, we know that by means of human speech higher-order conditioning can go to almost any length. In its most basic sense, language is nothing but a succession of paired presentations of stimuli, *i.e.*, a succession of sentences composed of paired "subjects" and "predicates." As we have already intimated (footnote 4), a word differs from the sounds and signs used in animal conditioning experiments principally in that words become highly abstracted, *i.e.*, able to convey their meanings without respect to the attendant physical circumstances. Their meanings may, of course, depend upon the *verbal context* in which they appear; but there are some words in which even this is not the case, *e.g.*, proper names.[9]

In other words, because words get reinforced in a highly generalized way (with respect to time and place), they take on a high degree of unconditionality; and this, we conjecture, is one reason that there is not the same limitation on higher-order conditioning by means of words at the human level that we find in the case of those signs used with animals. One may surmise that if signs were used and reinforced with an animal in a wide variety of circumstances, so as to give the animal explicit training in their generality (as we may indeed do, for example, in training a dog as a companion and pet), then the limitations encountered in the Pavlovian laboratory situation would be lessened. In other words, it is probably not just animal mentality that sets the limits on higher-order conditioning as Pavlov and others have studied it, but, perhaps equally, the limitations of the method employed.

In the main body of this paper, we have used the conventional terms for referring to the CS, UnCS, CR, and UnCR,[10] *i.e.*, we have used "conditioned" and "unconditioned" rather than "conditional" and "unconditional." But we may now ask how one might speak, more ideally, in respect to such

[9] For an interesting discussion of context and meaning, see Miller and Selfridge (9).

[10] That is to say, terms which are "conventional" within the framework of two-factor learning theory (11, 12). We have not used the term "conditioned response" to refer to a response, such as running, jumping, or lifting a paw, which is clearly a problem-solving response. In two-factor theory, conditioned responses are defined in a more restricted and precise fashion.

matters. From what has just been said it would seem to follow that the UnCS and UnCR may best be characterized as "unconditional." The CS and CR, on the other hand, can be most appropriately described as "conditioned" if there is a CS-CR connection, and if a stimulus is not yet conditioned or has been extinguished it might, in the one case, be referred to as "unconditioned" and, in the other, as "deconditioned." When a CS-CR connection is in force, it may be said to have varying degrees of conditionality, depending upon where it lies along the conditional-unconditional continuum as described above.

SUMMARY

It is the central thesis of this paper that in many experiments on conditioning this phenomenon is misconceived and mismeasured. We believe that conditioning, properly defined, often occurs very quickly and that conventional measures of this type of learning are often confounded with measures of *discrimination learning*.

If this inference be true, it follows that variation in the difficulty which subjects (rats) might be expected to have in distinguishing between the interstimulus interval (situation-with-CS) and the inter-trial interval (situation-without-CS) should produce marked differences in conventional measurements of "conditioning." This has been found to be empirically true.

The results of this study and their logical analysis further point up the desirability of differentiating *conditioning* and *problem-solving*, the former term denoting that type of learning which involves viscero-vascular tissue and the autonomic nervous system and the latter term denoting that type of learning which involves the striped musculature and the central nervous system. Discrimination is a phenomenon which can presumably occur with respect to either of these more basic forms of learning, but should not be confused with either of them. Said more precisely, discrimination learning represents an attempt on the part of living organisms to *resolve a conflict*; and while it can be largely derived from the more basic processes of conditioning (contiguity learning) and problem-solving (reward learning), it is a distinctive process, relatively complicated and slow of occurrence.

Finally it is pointed out that the debate as to whether one can most appropriately use the terms "conditio*ned*" and "unconditio*ned*" or "conditio*nal*" and "unconditio*nal*" to characterize the stimuli and the responses involved in conditioning is more than a semantic quibble; it involves the more basic problem of factoring out the extent to which the learning involved in any given situation is conditioning proper and to what extent it is discrimination learning. The conclusion is that the UnCS and the UnCR can be most appropriately qualified by the term "unconditional" rather than the term "unconditio*ned*," but that the CS and CR can best be qualified by the termed "conditioned," with "unconditioned" reserved for a stimulus which has not been or perhaps is to be conditioned and with "deconditioned" applied to a stimulus (or a response) after extinction has occurred.

REFERENCES

1. BROGDEN, W. J., LIPMAN, E. A., & CULLER, E. The role of incentive in conditioning and extinction. *Amer. J. Psychol.*, 1938, 51, 109–117.
2. CATTELL, R. B. *Personality: A systematic, theoretical and factual study.* New York: McGraw-Hill, 1950. P. 204.
3. COPPOCH, H., & MOWRER, O. H. Intertrial responses as 'rehearsal'—a note on 'over thinking' in animals. *Amer. J. Psychol.*, 1949, 60, 608–616.
4. GANTT, W. H. Psychosexuality in animals. In *Psychosexual development*

(P. Hoch and J. Zubin, Eds.). New York: Grune & Stratton, 1949. Pp. 33–51.

5. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century Co., 1943.
6. KORZYBSKI, A. *Science and sanity.* Lancaster, Pa.: Science Press, 1941.
7. LICHTENSTEIN, P. E. Studies of anxiety: I. The production of a feeding inhibition in dogs. *J. comp. physiol. Psychol.*, 1950, **43**, 16–29.
8. MCDOUGALL, W. *The energies of men.* New York: Oxford Univ. Press, 1938.
9. MILLER, G. A., & SELFRIDGE, J. A. Verbal context and the recall of meaningful material. *Amer. J. Psychol.*, 1950, **63**, 176–185.
10. MOWRER, O. H. Anxiety-reduction and learning. *J. exp. Psychol.*, 1940, **27**, 497–516.
11. ——. On the dual nature of learning—a reinterpretation of 'conditioning' and 'problem-solving.' *Harvard educ. Rev.*, 1947, **17**, 102–148.
12. ——. *Learning theory and personality dynamics.* New York: Ronald Press, 1950.
13. ——, & LAMOREAUX, R. R. Avoidance conditioning and signal duration—a study of secondary motivation and reward. *Psychol. Monogr.*, 1942, **54**, No. 5.
14. ——. Fear as an intervening variable in avoidance conditioning. *J. comp. Psychol.*, 1946, **39**, 29–50.
15. ——, & MILLER, N. E. A multi-purpose learning demonstration apparatus. *J. exp. Psychol.*, 1942, **31**, 163–171.
16. PAVLOV, I. P. *Conditioned reflexes.* (Trans. by G. V. Anrep.) London: Oxford Univ. Press, 1927.
17. ——. *Lectures on conditioned reflexes.* (Trans. by W. H. Gantt.) New York: International Publishers, 1928.
18. TOLMAN, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century Co., 1932.
19. WOLFLE, H. M. Time factors in conditioning finger-withdrawal. *J. gen. Psychol.*, 1930, **4**, 372–378.
20. WOODWORTH, R. S. *Psychology.* (Rev. Ed.) New York: Holt, 1929.
21. YARBOROUGH, J. U. The influence of the time interval upon the rate of learning in the white rat. *Psychol. Monogr.*, 1921, **30**, No. 135.

[MS. received August 12, 1950]

IDIODYNAMICS IN PERSONALITY THEORY WITH SPECIAL REFERENCE TO PROJECTIVE METHODS [1]

BY SAUL ROSENZWEIG

Departments of Psychology and Neuropsychiatry, Washington University, St. Louis, Mo.

Though, as an eminent psychologist recently remarked, "New projective methods are a dime a dozen," the conceptualization of these techniques still remains a scarce commodity. Only an occasional writer (4, 6) has recognized the conceptual difficulties and suggested possible solutions. The reasons for this lack are not far to seek. The projective methods of psychodiagnosis present unprecedented problems in psychological measurement, and these problems, in turn, arise from certain new conceptions of personality and of psychodynamics. Until these latter concepts have been systematically absorbed, the principles of test construction and standardization applicable to the projective techniques cannot be formulated.

I. The Individual and Norms in Psychological Theory

Part, if not the central aspect, of the new orientation in personality theory concerns the fate of the individual in the psychologist's perspective.[2] Elsewhere (9) an attempt has been made to delineate three stages in the history of psychology: the first, in which the individual has figured as the representative of the generalized human being to whom universal principles of segmental behavior could be applied; the second, in which he has been viewed as a more or less indivisible component of a statistically conceived group; the third, in which he emerges as a world of events constituting a population subject to both statistical analysis and dynamic conceptualization.

Closely tied to these conceptual destinies are three species of norms typically employed in psychological measurement and interpretation. Corresponding to the search for general laws of segmental behavior are certain implicit *universal* norms; in differential psychology—the psychometric approach—*group* norms on a statistical basis are explicitly employed; where, finally, the person in his intra-individual dynamics is taken into account, *individual* norms are recognized as crucial.

It is comparatively easy to illustrate the three kinds of norms by following the association method through its historical vicissitudes. After Galton invented it, Wundt and his students used the method to derive certain general laws of stimulus-response relationship in the tradition of the generalized human being and of universal norms. In the descriptive psychiatric province the technique was developed for purposes of diagnostic differentiation, *e.g.*, the work of Kent and Rosanoff, with group norms as the objective. With the advent of psychoanalysis and Jung's attempt to validate its concepts experimentally, the word-association method came to be employed as a probe for

[1] Presented in a preliminary version at the 1950 meeting of the American Orthopsychiatric Association.

[2] The obvious influence of G. W. Allport (*Personality*. New York: Holt, 1937), who has so effectively stressed the rôle of the individual in the science of psychology, will be recognized by any informed reader. The emphasis in the present account is, however, explicitly psychodynamic.

THE INDIVIDUAL AND NORMS IN PSYCHOLOGICAL THEORY

Theoretical Matrix	General—Experimental	Statistical—Psychometric	Psychoanalytic—Projective (Idiodynamic)
Objective	To establish principles of stimulus-response relationship applicable to the generalized human mind.	To allocate the individual in the group according to the distribution of various individual differences.	To reveal intra-individual (response-response or need-response) relationships for understanding the given case and for deriving principles of personality functioning.
Rôle of the Individual	An exemplar of responses in segmental stimulus-response behavior.	A constituent part or component of the group, the group being the basic unit.	A universe of events (idioverse) subject to analysis as a population.
Species of Norms	*Universal:* functional criteria of behavior, usually implicit, conceived as operating, under the required conditions, without exception for all individuals.	*Group:* statistical standards, usually explicit, for measuring the individual's conformity to the average of the group or for placing him otherwise in the range of its distribution.	*Individual:* implicit or explicit terms of reference for understanding the behavior of the person in the context of what is characteristic or idiomatic of him.
Criterion of Normality	Frequency of the conditions that initiate lawful sequences of behavior, the infrequent being the pathological.	Frequency of end-product behavior in a given social group or culture, the infrequent being the abnormal.	Economy or efficiency in intra-individual organization, the uneconomical being the unhealthy.

individual norms, called by Jung "complexes." The survey indicates that a single psychological tool can be employed with quite different objectives as regards the rôle of the individual and related norms. Furthermore, this example may serve to demonstrate that all projective methods, despite their indigenous bias towards individual norms, allow as well for group and universal norms in a complete conceptualization.

The accompanying chart summarizes the varying rôles of the individual and the attendant, typical norms encountered in the history of psychology and in contemporary psychological theory. As has already been implied, and as will appear later in further detail, the objective of each theoretical matrix is different and in each the individual assumes an appropriately modified conceptual rôle. The norms that implicitly or explicitly prevail in the making of interpretations vary similarly and, as the chart indicates, the criteria of normality or abnormality conform to pattern in each approach. The intended implication of the survey is that, while it is theoretically possible to take consistently different views of the individual, the current trend is toward an idiodynamic formulation in which the person is studied with the dual purpose of understanding him as an individual and of deriving general principles of intra-individual functioning. Despite this emerging emphasis, the other two orientations are by no means obsolete and must be taken seriously into account.

II. IDIODYNAMIC PERSONALITY POSTULATES

To appreciate how the projective methods can combine the various concepts of the individual, full recognition must first be accorded the "idiodynamic" approach to personality. The projective techniques embody this approach in at least three closely related aspects: response dominance, configuration dominance, and idioverse dominance. These postulates are not only implicit in the methods of psychodiagnosis but are rooted in psychodynamic theory generally.

1. *Response Dominance.* Experimental psychology, as well as the psychometric clinical approach, has always proceeded on the assumption that stimulus-response relationships provide the key to the understanding of behavior. In order to build a science of psychology the stimulus had to be a public matter, externally defined, whether by the

experimenter or by the examiner. The formulation of behavior principles in which such externally defined stimuli were coördinated to certain regularly expected responses constituted the goal of the traditional experimentalist. Similarly the psychometric examiner effected his measurements by relating the responses of his subject to certain standardized stimuli for which there were pre-defined expectancies of response. With the introduction of the psychodynamic point of view came a revolution in the conception of the stimulus and of its relation to the response. The stimulus was dethroned; the dominance of the response ensued. The relationship of responses to each other in a given individual was now the matter of primary interest.

To the traditional psychologist this new state of affairs represented a decided threat to the objective nature of science. In the new regime he saw the degeneration of psychology into a subjective, impressionistic game whose proponents were at best clinicians and at worst charlatans. To the psychodynamic theoretician and to the projective diagnostician, however, the new approach to the stimulus was imperative since it thus became possible to accord the individual personality its proper place in psychological study and theory. What is the answer? The difference between the older, traditional approach and the newer, psychodynamic one is not a matter of objectivity but of objective. The psychophysicist, the behaviorist and the psychometrician assume certain constant stimulus-response relationships and operate accordingly; the psychodynamic theoretician is interested in the response-response relationships to be found in the given individual. The change is from a search for stimulus-response relationships to the study of response-response relationships. Concurrently the generalized human being yields place to the unique individual in central interest for the psychologist.

Thurstone long ago recognized this shift and discussed it under the heading "stimulus-response fallacy" (14). Writing in 1924, he appealed in his exposition to the "New Psychology"—psychoanalysis—and credited it with cogently emphasizing relationships between the impulse or need and the response of the individual, in contrast with the more academic practice of concentrating upon relationships of stimulus and response. In the light of this statement the postulate of response dominance may be more sharply defined. Response-response relationships are significant not only in their own frame of reference but because they also reflect need-response relationships. The translation of manifest to latent content in Freudian dream analysis and the similar interpretations in the projective methods illustrate this type of inference. The rôle of the external stimulus may also now be clarified. In the formulations which stress response-response or need-response relationships, the stimulus is often merely a trigger, sometimes only an "excuse," for evoking prepared reactions that have already been set in motion by the inner stimulation of an impulse. It is on this basis that the external stimulus is frequently regarded as being *selected* by the individual or by his impulse-nature from a multiplicity of possible external stimuli. Past stimuli, in so far as they have modified the inner organization of the personality, may thus be far more influential in determining present behavior than any environmental stimulus acting at the moment. Rôles are thus assigned to need or impulse, inner stimulation, past stimuli, present external stimulus, and response.

To the projective methods the postulate of response dominance is applicable with one important qualification. While

in the free-association method of the psychoanalyst no regard for a standardized stimulus of any kind is necessary, the concern being exclusively with relationships between the subject's earlier and later associations, in the field of the projective techniques there are usually certain common stimuli that are employed consistently with all subjects. One of the advantages of the projective method lies precisely in this qualification. But this difference between the projective techniques and other psychodynamic methods of study complicates the problem of measurement and demands the recognition not only of individual norms but of group norms—stimulus-oriented—as well. The difference between the word-association test as a projective method and the free-association method as the psychoanalyst employs it illustrates the point.

Another way of recognizing the qualified application of response dominance to the projective techniques is to observe that the examiner presents his standard stimuli primarily as a means of eliciting responses. The interrelationships of these responses contain the essence of what he is seeking—the organization of the so-called "inner personality." The stimulus as such is more or less negligible—*more* in so far as the unique responses of the individual subject delineate him psychodiagnostically, *less* in so far as the stimulus itself is a standard one applied to all subjects and related, within certain limits, to group-standardized responses that permit a certain amount of prediction about the subject's relationships to other subjects. The nub of the problem again lies in the interplay of individual and group norms and their mutual definition.

It should further be noted that universal norms are by no means ruled out in the psychodynamic approach. Concepts like repression and displacement are obviously considered to have a universal application under the appropriate conditions. But there is a difference between such general principles and the corresponding behavioristic stimulus-response formulations. In the last analysis the conditions for repression and the threat to the ego which makes for the operation of this form of defense are to be found in the individual's definition of the stimulus situation according to his own past experience and present attitudes. Moreover the principle itself is mainly a statement of intra-individual dynamics rather than of any systematic relationship between externally defined stimulus and internally mustered response. For the very distinction between external and internal has largely disappeared and is gradually being replaced by a phenomenological frame of reference in which the "stimulus" and the "response" are interdependent aspects of the subject's universe. Since the stimulus is regarded as ultimately defined by the response, the former can no longer serve as an infallible determiner of the latter.

2. *Configuration Dominance.* The twilight of the stimulus is matched by the similar subordination of the part to the whole in the idiodynamic approach. The phenomenological relationship between stimulus and response is paralleled by the absorption of parts into wholes. The important influence of Gestalt psychology is here recognized as complementing the already indicated influence of psychoanalysis. The wholeness of the subject's behavior lies both in the pattern of his responses and in that of the standard stimuli projectively employed. The essential totality is, of course, the world of the individual in which "responses" and "stimuli" are mutually definitive.

Stated thus baldly the postulate of configuration dominance is as much a threat to the traditional psychologist as is the postulate of response dominance. Both signalize the eclipse of scientific "objectivity." Increasingly, however,

empirical work has demonstrated both in Gestalt psychology and in experimental psychoanalysis, if not as yet in the field of the projective methods, that these postulates are denotatively stable, not merely a confession of faith in the uniqueness of the individual. A new conception of objectivity is perhaps implicit in recent approaches to the person, but the fact that the new objectivity differs from the old is in itself no criterion of adequacy. It must in the end be recognized that the conception of objectivity inherited from nineteenth century science rests upon a one-sided, if not prejudiced, view of the individual and his place in nature. The final test must, as always in scientific method, rest with the evidence. When it becomes possible in the course of time and experience to compare the effectiveness of the older and newer approaches in terms of their capacity to allow prediction and control, the relative merits of each can be determined.

An important aspect of the part-whole problem as applied to psychodynamics and the projective methods is found in the conception of personality levels (7). In attempting to understand the individual as such, it is essential to recognize various levels or layers of the personality, either in terms of conscious and unconscious, as id, ego and super-ego, or, perhaps more demonstrably, as opinion, overt and implicit levels of response. In whatever terms the formulation is made, it becomes incumbent upon those working with projective methods and psychodynamic concepts to distinguish the interplay of personality levels in any observed behavior. Interpretation, as the concept is employed in psychodynamics, rests largely upon this distinction, *e.g.*, in Freud's differentiation of manifest and latent dream content or in Lewin's distinction between phenotypical and genotypical levels of explanation.

3. *Idioverse Dominance.* This third postulate has been partially anticipated and exemplified in the previous argument. In fact, all three of the postulates here outlined are aspects of a single reorientation. But each aspect deserves separate emphasis.

Idioverse is the name here given to an *individual's universe of events.* The concept is implicit in psychoanalytic theory and has been advanced in more than one independent version (1, 11, 13); the term is new. Idioverse dominance highlights the standpoint of the individual in any attempt to understand him. At the same time, as has already been noted, the individual is conceived not to exist in a world completely separated from the worlds of others. Thus individual norms, *e.g.*, complexes, play an important part in the assessment of personality but, at the same time, group norms, in terms of shared experiences of a common stimulus, also play a part. Universal norms and general principles based upon them contribute to idiodynamic theory, though—to repeat—these general principles differ from the traditional ones of experimental psychology in the extent to which intra-individual organization is emphasized and in the way in which the stimulus is defined "subjectively" by the individual's own performance. The inclusion of group and universal norms, as well as individual ones, is wholly compatible with idioverse dominance. Frequently the error is made both by opponents and proponents of the projective techniques of identifying these tools with the approach of individual norms alone.

It follows from the postulate of idioverse dominance that the projective technique is not to be conceived as a "test" in the usual sense.[3] The test,

[3] If the reader notes the sedulous avoidance of "test" as a substantive following "projective" and a consistent preference for "method," "technique" or "procedure," the reason is ap-

as it has become known in the field of psychometrics, exists for the measurement of individual differences in a group, the norms being statistical or group norms. To place the individual in the group becomes the aim of the procedure. External standards, preestablished by the examiner or by his colleagues who standardized the test, determine right and wrong responses—just as certain expected responses to objectively defined stimuli figure in the principles of behavioristic, general psychology. In the projective technique, as contrasted with the psychometric test, the approach is shifted from an emphasis upon group norms to an emphasis upon individual norms—the more or less unique ways in which the subject structures or interprets the eliciting stimuli. As has been pointed out elsewhere (10), however, it is still necessary to employ group norms, wherever feasible, in order to demarcate the individual from the group. If, for example, one does not know what is commonly given as a response to a particular TAT card, it becomes impossible to say what is unique. Group norms thus play a part in the detection of individual responses and of the individual norms implicit in them. Moreover the delineation of individual norms serves not only to show how the person functions in his particular universe but also to prefigure the inner organization of his personality according to the relatively universal principles of psychodynamics. The universal norms of psychodynamics thus combine with the individual norms adumbrated in the projective methods to prepare the way for laws applicable to the population of events in the individual universe.

The conceptualization of the projective techniques is possible only on the basis of the foregoing postulates. Response dominance, configuration dominance, and idioverse dominance involve a complete revolution of thinking in psychological measurement. The familiar concepts of test construction and of experimental design are obviously incompatible with such a radical reorientation even though, as in all revolutions, the past is to a certain extent implicit in and serves as an evolutionary foundation for the over-emphasized present.

III. The Conceptualization of Projective Techniques

From the foregoing exposition of personality theory as related to the projective methods it should be easier to appreciate the difficulties which these techniques present, as compared with the usual psychometric test, when problems of test construction and standardization are posed. The present discussion of these difficulties will be limited to the two chief dimensions in respect to which the traditional test has had to prove itself. These dimensions are, of course, *reliability* and *validity*.

1. *Reliability*. The inapplicability of the traditional reliability concept to projective techniques can be readily appreciated if the previous discussion of configuration dominance is recalled. The standard stimuli in the projective method are almost always considered to constitute a configuration that cannot be analyzed into its separate parts or segments without significant loss of meaning. Thus the Rorschach cards have certain relationships to each other on the basis of which it is possible to detect so-called "color shock." Changes in the reaction time from card to card and differences in the quality of re-

parent at this point. The projective *method* is not to be confused with the psychometric *test*. The term "test" has been preëmpted for use in the orientation of the stimulus, the segment, and the external universe of the person as something imposed upon the subject from without. To respect this usage promotes clarity.

sponse, all of which derive from the fact that the subject has first seen certain cards and is then shown others, depend upon the configurational unity of the total series. Once this fact is recognized it becomes obvious that the reliability of a projective technique cannot be determined by either the split-half or alternate-item methods. Not only are the parts or aspects of these psychodiagnostic instruments interdependent, but even when separable they differ from the parts of a psychometric test in being of unequal psychological import. This inequality is, moreover, not merely a quantitative but is rather a qualitative matter. The inequality of units refers not to differences in the difficulty of items or in the capacity of the subject to solve them, but to differences in the kind of psychological reaction the stimuli may be expected to elicit. Under these circumstances any attempt to divide the total series of cards into a first half and a second half or to equate odd and even items for the purpose of determining reliability would violate the essence of the procedure. Any two "halves" of the method are simply not intended to be equal and hence should not correlate with each other if the projective technique lives up to its configurational design.

One might then maintain that it should at least be possible to use alternate forms of a projective technique and establish reliability by this means. To a certain extent the suggestion has been borne out in practice. Thus one finds for the Rorschach method not one but two alternate forms—the Behn and the Harrower. In theory the solution is less simple. If one is to examine the same subject on two different occasions with either the same or alternate versions of a particular projective technique, it is still necessary to consider the possibility that the experience of participation on the first occasion may affect the second performance. The temporal configuration of the subject's total experience here manifests itself in a new form. The nearest equivalent to this dilemma in the psychometric test of capacity is the effect of practice on performance, but since ability is presumably static over a brief period of time, such effects are negligible or can be statistically corrected. In the projective methods such solutions are not available, for one is there attempting to evaluate subtle personality processes with a temporal dimension that includes the configurational effect of earlier upon later experience.

Moreover any theory of personality must take into account the general fact that a person is not identically the same from the time of one experimental or psychodiagnostic session to that of another. To be sure, there are presumably stable aspects of personality which have a certain degree of permanence and should be measurably equal from one day to the next. However, these stable and basic aspects are not yet easy to define or to tap. Mood changes are sometimes noted to affect the results even of the venerable Rorschach. In other words, to study the subject's performance on several occasions is not merely to assess the reliability or stability of an instrument—as is true for the psychometric test; one is inevitably committed to the measurement of personality consistency. The familiar reliability concept is, therefore, too narrow for the projective field.

In some of the writer's experimental work the foregoing considerations as to the repeatability of a subject's performance have led to a novel concept—"the range of personal variability." The concept is intended to allow for the scope which any personality must be permitted if it is to maintain its constantly changing and dynamic character. On the assumption that personal-

ity is constituted of a population of events (the idioverse), the range of variability here in question corresponds to the standard deviation or some other measure of dispersion around the mean performance. A statistical analysis of the intra-individual population is thus envisaged as part of any reliability determination. It is hoped that by concepts of this kind it will be possible eventually to adapt or replace the traditional techniques of test construction and standardization. In such efforts the theory of personality and of psychodynamics will, of course, have to be closely coördinated with the statistical methods evolved.

In the meantime it is well to bear in mind that the reliability dimension has its own limitations. Reliability as such is no unmixed blessing, especially in the field of personality assessment. It is self-evident that a high degree of retest reliability, demonstrated for a particular personality inventory, may have been achieved by virtue of an equal degree of invalidity on both occasions. If the subject commits the same "opinion errors" both times, the scores of his tests will, of course, correlate highly. The important conclusion will, however, not then be that the inventory has proved to be reliable but rather that it has proved—if we can prove it—to be equally *invalid* on the two occasions. Validity thus becomes a far more basic matter than reliability—a point to which it will be necessary to return.

A confusingly related consideration is the reliability of scoring. In psychometric tests, *e.g.*, the Stanford-Binet, response samples and other means of making it possible for different examiners to score identically have played an obvious part. When, however, a projective technique is under scrutiny, the reliability of scoring becomes more problematic, since here the variety of possible responses is so much greater and their anchorage to the stimulus is so much looser. In this context Harrison and Rotter (5) have attempted to show that the TAT can be consistently scored by two independent workers. A similar project has been devoted to the Picture-Frustration Study (2). Reliability of scoring as here recognized, should not, however, be confused—as Harrison and Rotter tend to do—with test reliability in the usual sense.

2. *Validity.* Far more intricate than the problems associated with reliability are those connected with validity. As Ellis (3) has shown in his critical review, the validity of personality inventories is notoriously weak as compared with their reliability. In fact, it almost appears as if the student of personality has consoled himself with high reliability coefficients in compensation for any adequate degree of validity. In contrast to this state of affairs it should be noted that proponents of the projective methods have always put validity ahead of reliability. Their demonstrations of validity have not, however, always convinced others. Clinicians are apt to be too easily satisfied with the consistency of their own experience as a proof of validity. Some of them even make a fetish of the fact that the methods they use do not need, or at least cannot benefit from, the time-honored techniques for establishing validity. Anyone who has had clinical experience in the appraisal of personality by projective and cognate methods can testify to this temptation.

The student of personality theory and measurement cannot afford to be so sanguine. He requires systematic evidence, carefully controlled. At the same time he takes issue with those investigators from the psychometric field who, not being conversant with the psychodynamic basis of the projective techniques, tend to ignore the postulates reviewed above. The most com-

mon error of this kind involves the use of segmental "objective" methods of determining both reliability and validity. Vernon (15) early recognized the limitations of such approaches to the Rorschach and offered the matching method as one corrective.

In keeping with this orientation the first aspect of the validity problem in the projective methods concerns their configurational properties. To be completely consistent one might even question whether the scoring of projective productions, segment by segment, does not violate their Gestalt character. The insistent need for analytic scoring if one is to come to grips with the data has denied this point the systematic attention it deserves. Perhaps this partial blindness implies a recognition that no alternative exists except a frankly impressionistic and suspect approach to the subject's total productions. But before the alternative is dismissed, it might be useful to consider with Sarason (12) the possibility of putting so-called *subjectivity* or impressionistic interpretation (as opposed to scoring) on an objectively phenomenological basis. There is nothing to prevent empirical studies of the comparative validity of impressionistic interpretations with others arrived at by the more usual analytic-synthetic procedures.

Closely related to the scoring issue is the problem of interpreting particular quantities or values of an isolated scoring factor. Beck, for example, has repeatedly emphasized that identical quantitative values for any one of the Rorschach factors may have quite different psychological meanings depending upon the total pattern of scores in which that quantity is embedded. If, therefore, one is to validate a projective technique, it is necessary theoretically to consider this multipotentiality of the isolated factor. One must take into account the shifting meaning of its whole range of quantities, each one of which can, in turn, have a multiplicity of meanings according to whether it is found in any one of a multiplicity of scoring patterns. As can readily be seen from this rather appalling, but not exaggerated, statement of the case, the prospect involves an almost limitless number of situations to be validated.

But it is necessary to validate more than the individual factors of a projective technique; the various patterns of such factors lie even closer to the configurational approach which these methods espouse. According to the previous exposition of idiodynamic postulates, it could even be maintained that individual scores or measures have a partial, if not misleading, significance psychologically and that only the total record or protocol should be interpreted. Some hope is perhaps to be found in following this line of thought to its logical conclusion. Are there possibly certain qualitatively unitary aspects of personality—certain qualitative quanta—that are more or less finite in number and that may be isolable in the various patterns that a projective technique discloses? If such a hypothesis is justified, the aim of validating various separate quantities of a particular scoring category would be abandoned. Only configurations of scores would be considered as having possible significance, and a subject's record would be scrutinized as a whole or as a constellation of wholes with certain possible unitary meanings.

Having considered the special properties of the projective technique as they affect validation, it is necessary to examine next the broader issues of the validation procedure in its entirety. The first of these issues is the definition of the validity objective. To ask in general terms whether a projective method is valid or not is relatively meaningless. It is always essential to

decide for what particular purpose—for what specific predictive or heuristic use—the validity of the procedure is to be assessed. Far too often this obvious caution is not observed, and the results obtained are then bound to be confusing and inconclusive.

The definition of the independent validity criterion is another frequent source of confusion. In particular, the problem of levels of personality expression (7) is here noteworthy. Whether the subject is manifesting the opinion (self-critical), overt, or implicit level of behavior, or some combination of them, is not easy to determine. Yet if a projective method is tapping one level, say the implicit, and the judgment of personality with which the technique is to be correlated for purposes of validation is made at another level, say the overt, the fact that little agreement between these two assessments were found to exist would in no way reflect upon the validity of the instrument. Unless this difference in levels had been taken into account, false conclusions might be drawn regarding the validity of the technique.

The problem of validating the independent criterion is a further aspect of its proper definition. When one resorts to an independent criterion, one turns inevitably to one or another of three main sorts of evidence: one may make or ask for certain observations of the subject in situations independent of the projective method; one may obtain ratings from some judge who has observed the subject under favorable conditions; or one may apply other measuring instruments with the intention of correlating their results with those of the instrument to be validated. The hazards connected with each of these courts of appeal are readily apparent and have been repeatedly described. Observations made on a subject are inevitably colored by the situation in which he was observed or by the personality of the observer, and such evidence is therefore often as dubious as the evidence of the projective method to be validated. Ratings are similarly apt to be subjectively contaminated, even when the judge is highly motivated and conscientious in the performance of his task. Measures from other instruments, especially if projective, are themselves apt to be questionably valid. To validate one unvalidated projective technique by means of another involves the investigator in a circular type of argument, with the blind leading the blind. It may be concluded that no investigation of validity can be more than suggestive or preliminary unless precautions are taken in it to establish the validity of the independent criterion.

To accomplish the above goals in respect to validation new investigative procedures may be indicated. One example is the method of validation by successive clinical predictions (8).

3. *Relative Values of Reliability and Validity.* Up to this point reliability and validity have been considered more or less separately. It may now be useful to cast a glance at their interrelated significance. As has been indicated above, reliability without validity can be a doubtful virtue. If one must make a choice, it is obviously validity that is fundamental.

It can be convincingly maintained that validity demonstrated on repeated occasions implies reliability. To put the matter in these terms has a certain advantage from the standpoint of personality theory since stress is here placed upon the importance of measuring what one purports to measure. Reliability as such is more a property of an instrument than it is a reflection of the relationship between an instrument and the personality one seeks to assess. The only merit in getting the same answer twice with a particular measuring

rod lies in the fact that the measure actually reflects some property of that which is to be measured. In so far as personality is itself inconsistent or subject to change with time, room is rightly left for a certain degree of psychometric "unreliability" without any contradiction to the validity of a projective technique or to its *essential* reliability.

IV. SUMMARY

In the history of psychology the theoretical fate of the individual and the attendant norms for interpreting his behavior have evolved in an idiodynamic direction. The essential postulates of the idiodynamic approach are response dominance, configuration dominance, and idioverse dominance. Personality theory of this type is fundamental to any compelling conceptualization of the projective techniques, but cannot be used in isolation from the relevant aspects of the general-experimental and psychometric points of view. In thus integrating personality theory with the newer methods of psychodiagnosis, the traditional concepts and methods of test construction and standardization, notably reliability and validity, must be recast.

REFERENCES

1. ALLPORT, F. H. Teleonomic description in the study of personality. *Character & Pers.*, 1937, 5, 202–214.
2. CLARKE, H. J., ROSENZWEIG, S., & FLEMING, E. E. The reliability of the scoring of the Rosenzweig Picture-Frustration Study. *J. clin. Psychol.*, 1947, 4, 364–370.
3. ELLIS, A. The validity of personality questionnaires. *Psychol. Bull.*, 1946, 43, 385–440.
4. FRANK, L. K. *Projective methods.* New York: Thomas, 1948.
5. HARRISON, R., & ROTTER, J. B. A note on the reliability of the Thematic Apperception Test. *J. abnorm. soc. Psychol.*, 1945, 40, 97–99.
6. MACFARLANE, J. W. Problems of validation inherent in projective methods. *Amer. J. Orthopsychiat.*, 1942, 12, 405–410.
7. ROSENZWEIG, S. Levels of behavior in psychodiagnosis with special reference to the Picture-Frustration Study. *Amer. J. Orthopsychiat.*, 1950, 20, 63–72.
8. ——. A method of validation by successive clinical predictions. *J. abnorm. soc. Psychol.*, 1950, 45, 507–509.
9. ——. Norms and the individual in the psychologist's perspective. In *Feelings and emotions* (M. L. Reymert, Ed.). New York: McGraw-Hill, 1950.
10. ——. Apperceptive norms for the Thematic Apperception Test. I. The problem of norms in the projective methods. *J. Personality*, 1949, 17, 475–482.
11. ——. Projective techniques: their progress in the application of psychodynamics. In *Orthopsychiatry 1923–1948: Retrospect and prospect.* New York: Amer. Orthopsychiatric Assoc., 1948. Pp. 456–469.
12. SARASON, S. B. The TAT and subjective interpretation. *J. consult. Psychol.*, 1948, 12, 285–299.
13. SNYGG, D., & COMBS, A. N. *Individual behavior.* New York: Harper, 1949.
14. THURSTONE, L. L. *The nature of intelligence.* New York: Harcourt, Brace, 1924.
15. VERNON, P. E. The significance of the Rorschach test. *Brit. J. med. Psychol.*, 1935, 15, 199–217.

[MS. received September 1, 1950]

SOME THEORETICAL CONSIDERATIONS OF LEARNING[1]

BY W. J. BROGDEN

University of Wisconsin

I am not opposed to theories of learning nor to their development. It is my opinion that some phenomenon of learning is involved, either directly or indirectly, in the conduct of almost every psychological experiment. If the data collected are not the acquisition, retention, or transfer of performances by subjects, they are dependent upon prior acquisitions and their retention or transfer. Thus learning is the core of the science of psychology. The advancement of knowledge in other areas of psychology depends upon the advancement of our understanding of learning. General laws of learning, then, are the goal. Although laws in the last analysis have their origin in experimental results, theory development can assist materially in such progression. A theory may organize the results of many researches, it may bring new relations to light, or it may serve as a catalyst for fruitful experimentation. On the other hand, a theory of learning may impede advancement seriously. It may fail to consider existing experimental evidence that does not support it; it may encourage research to proceed in non-productive channels; or it may define problems verbally that cannot be attacked experimentally. There are other ways in which theories can hinder advancement or accelerate it. Each of our present-day theories is probably doing both.

It is not my purpose, however, to examine present-day theories of learning separately, nor to criticize some and support other parts of their structure. Rather I hope to comment constructively on some general problems that I believe are common to the development of theories of learning.

The first of these problems is that of the units for describing psychological phenomena. I shall assert that all psychological phenomena are samples of behavior—reactions, responses, or ways of conduct. This is a well-beaten path, and therefore not difficult to follow, but the next step, that of classifying all of the different kinds of behavior, is the crux of our taxonomic problem. There has been no successful attempt to itemize ways of behaving and to locate each item within a comprehensive classification system. One finds only the rough general classes that are made by the adjectival modification of the nouns *behavior* and *response:* learned behavior, unlearned behavior, intelligent behavior, emotional behavior, unconditioned responses, conditioned responses, verbal responses, motor responses and the like. There have been proposals that behavior be classified in terms of the stimuli that evoke it, or in terms of the goals toward which it is directed. No systems of classification have been developed from these suggestions that are adequate for either experimentation or for the construction of theory. It probably is not possible to achieve a comprehensive classification of behavior at the present time. However, experimentation and theory construction will be hampered, and the formulation of general laws will be impossible until this has been done. I wish to propose several steps that should aid in the solution of this problem.

One is to limit the frame of reference for theories of learning to experimental

[1] Presidential Address, Division of Experimental Psychology, State College, Pennsylvania, September, 1950.

results and the laboratory conditions under which the results are obtained. Empirical observations of one's dog, one's children, one's fellows, one's client or one's self have no place in the construction of theory, nor is their use legitimate in arguments that purport to favor or oppose a given theory. A scientific theory is relevant only to the facts and relations obtained by scientific methods. It may account for incidental observations made in everyday life, but it cannot be held responsible if it fails to do so. If theories of learning are constructed to encompass only the experimental study of learning, and are treated only in this context, both development of theory and experimentation will contribute more effectively to the solution of the classification problem and to the formulation of general laws of learning.

Another forward step can be made by standardization of apparatus and of methods and procedures used in experiments on learning. This is easier said than done. It has been said before, and we still have marked variation in many of the conditions of experiments that deal with the same problem, even in the case of a series of experiments completed by a single research worker. The only suggestion that I can make is that the Division take responsibility for solution of the problem of standardization, not only for learning, but for all areas of experimental psychology. It is obvious that standardization would accomplish much. Not the least of these accomplishments would be that of providing an excellent basis for a system of classifying behavior.

The major step towards solution of the taxonomic problem would be taken if there were available for each of our psychological laboratories a pool of subjects whose lifetime job was that of being a subject in psychological experiments. Each subject pool would have to meet many conditions and there would be special requirements for any given pool in order that it be of maximum utility to a particular laboratory. The obstacles to be overcome in organizing such a supply of subjects are enormous, but I am convinced that at some time in the future something of this sort must be done if the frontiers of psychology are to be advanced further. Experimental case histories on subjects would be a tool of tremendous value. Think of what this would mean in solving the problem of classifying behavior, in the design and conduct of experiments, and in the development of theories. We should then be able to find out whether or not a psychology based on studies of the college sophomore and the white rat differs from a psychology based on studies of human beings. Obviously the only possible financial source for such a program is the Federal Government. With the National Science Foundation now coming into being, we should seek support from it for such a program.

A second general problem for theories of learning has to do with whether or not there is more than one type of learning. This also is a problem of classification. It may or may not be the same as the problem of classifying behavior. Any apparent differences in types of learning may be no more than differences that exist between different classes of behavior. If this is so, the problem will be solved when an adequate taxonomy of behavior has been developed. On the other hand, some theories of learning distinguish between two or more types of learning, and this proposal can certainly be given consideration. Learning is not a well-defined phenomenon. It covers a very wide range of events and conditions, and it is not always easy to distinguish phenomena of learning from other psychological phenomena. Although learning

includes acquisition, transfer, and retention, in the context of learning theories it is almost always synonymous with acquisition. It is in this sense that the term learning will be used in the following discussion.

Learning is a change in the behavior of an organism. The change has direction from an initial quantity or proficiency of performance to a different quantity or greater proficiency of performance. Its course is limited by the arbitrary initiation and termination of measurement. It is a function of the repetitive or continuous presentation of the conditions under which measurement of the change in performance is made. These are the salient characteristics of learning as it is measured experimentally. Differences in learning occur in terms of each of these characteristics and in their combinations. Distinctions between types of learning occur primarily in terms of the conditions under which measurement of the change in performance is made. But not all theorists are willing to deal with learning in terms of performance. Some insist on a distinction between the two. Learning is not performance, but it is inferred from performance. The inferences are called intervening variables or hypothetical constructs, and it is in these terms that learning is described and defined.

It is not necessary that learning be inferred from performance. Learning can be measured as directly as any item of behavior. It has a beginning and an end, and this gives it the characteristic of duration. Since it is a change in performance with direction, it has a latency and an amplitude. Its course is given by the learning curve. It is true that learning is measured in terms of some property of an item of performance, its latency, amplitude, relative frequency, or rate. But this does not mean that learning is something other than performance, and that it can only be inferred from performance. It demonstrates that learning is performance, and that the direct quantification of learning is possible only in terms of performance. Since the intervening variable has status only as inference, it cannot be dealt with directly in experimentation. Its existence lies in the realm of theory and its evaluation as a tool for theory will depend upon evaluation of theories that make use of it.

The distinction of different types of learning will depend in part upon how learning is measured and whether it is treated directly as performance or as an inference from performance. But the primary distinction of different types of learning, if they do exist, will be in terms of conditions. The range of conditions under which learning occurs is extensive. New conditions of learning are being uncovered by experimentation all the time. This situation will undoubtedly continue indefinitely. In fact, there may be no finite limits to the conditions that can influence learning. With such a wide range of conditions, and a generality of learning such that it extends into every area and field of psychology, distinctions between types of learning can readily be made. The theoretical problem is not whether there is more than one type of learning, but how many types there are. It is by means of the dimensional analysis of the conditions of learning that this problem will be solved. This cannot be done by experimentation alone. We need scales that will locate the various phenomena of learning on precisely defined dimensions, so that quantitative differences can be related to the various conditions of learning also expressed quantitatively. The determination of degrees of similarity, and of differences in the quality or quantity of learning, is our persistent headache.

Theory development should be able to contribute much here.

We have theories of learning in large part because of the assumption that for learning to occur there are certain conditions that are necessary. We have different theories of learning because different conditions are assigned this special value. Neither of these assumptions may be correct. The assumption of a limited set of conditions always necessary for the occurrence of learning is restricted to theories that do not differentiate between different types of learning. If we are able to differentiate between different kinds of learning, the assumption will be changed as follows: for learning of a particular type to occur there is a certain set of necessary conditions. Theories then would differ among themselves not only in terms of the assignment of special value to certain conditions for a given type of learning, but also in terms of the types of learning that they would admit. Arguments about theories tend to be arguments about the values assigned to certain conditions. If we develop theories that distinguish between types of learning, some of the disagreement between theories will disappear, since certain critical conditions will be assigned to one type of learning, and others to other kinds of learning. But as long as there are different theories of learning, their differences will be primarily in terms of the assignment of critical value to certain conditions.

All theories contain positive assertions, and some contain negative assertions. The positive assertions of one theory implicitly carry negative assertions relative to the positive assertions of another theory with which it is in disagreement. Negative assertions in a theory, whether explicit or implicit, represent the same order of logic as that of drawing conclusions from an experiment with negative results. If the results of an experiment are negative, there is no way by which one can assign cause to a particular variable or set of variables. It may be interesting to speculate about the conditions in a negative experiment, and such speculation can assist in redesign of the experiment so that it produces positive results. But speculation about the results of negative experiments sometimes intrudes into theories, and theories are then developed in terms of these speculations. If it is possible to draw negative conclusions, it can be done only from an exhaustive series of experiments in which all of the variables present in the experimental situation have been varied throughout their dimensions and in combination with each other. This, however, is not usually a productive type of research, and few experimenters have engaged in it.

Negative results in a learning experiment are always a relative matter. Usually only a single item of the organism's behavior is under measurement. Not all variables other than the independent variable are held constant, and some of these may vary systematically rather than in random fashion. There may be variables present that have not been identified, or that the experimenter considers have no significance for the experiment. Negative results may be negative only in the terms of the response designated by the experimenter, or they may be a function of interactions between variables that are involved in the experimental situation.

I had some serious difficulties with the first dog I tried to condition. I was shown how, I was told how, and I had read how. At the end of several training sessions the dog was not giving any conditioned responses, so at the next session I increased the intensity of the shock. There were still no conditioned responses, so I again increased the intensity of the unconditioned stimulus.

On a given trial, the animal struggled, and chewed through the strap that provided for attachment of a lever to the right foot, the foot that was being shocked and that was to be conditioned to the bell-conditioned stimulus. I got some leather, cut it, and riveted it together to make a new strap. When this was attached to the lever and to the dog's foot, I continued the training session. It was not many trials before the animal responded to the bell again by chewing the strap in two. Shortly thereafter my career as a psychologist was giving way to that of harnessmaker, and the dog had a stable conditioned response of chewing a strap in two at the sound of a bell. When I sought advice, it was pointed out to me that I was using a shock of too great intensity, the lever strap was too long, the head stock was too high, a collar clip was not being used, and the belly strap was adjusted too loosely. Too much shock produces struggling and diffuse activity. A long strap facilitates its being chewed in half. The adjustments of the head stock and belly strap force a posture on the dog from which forelimb flexion can be made only with great difficulty. With the appropriate changes in these experimental conditions, further application of the training procedure resulted in elimination of the conditioned chewing response and quick establishment of a stable, clearly defined conditioned flexion response to the sound of the bell.

If learning does not occur with the absence of a particular condition, it does not follow that this particular condition is necessary for learning. The negative results are in terms of learning as defined by the experimenter, and may be the result of conditions the presence of which the experimenter is unaware or the importance of which he is unable to recognize. Changes in the quantitative combination of conditions under which negative results were obtained may produce learning. The introduction of a different variable into the combination of conditions may also produce learning. A theory of learning that assigns a critical value to a given set of conditions states in effect that learning will not occur in the absence of these conditions. It seems to me that it is more fruitful to assume that learning of some kind always occurs in experiments conducted on normally receptive and responsive organisms. Thus if negative results are obtained, they are recognized as being negative only in terms of the learning problem set by the experimenter. The immediate problem, then, is one of finding what kind of learning did occur in the experimental situation, and the later problem to differentiate the conditions that produce the acquisitions that are in competition. For theory development this assumption would provide the same problem at a more general level and the critical value of variables would be assigned to this learning or that learning, not to this learning or no learning.

There is a relation between theory and experiment in the field of learning that I consider to be a detriment to both. The so-called crucial experiment has its origin in theoretical formulations of learning that assign critical values to certain conditions. It is a kind of experiment that is simple in design and complex in its theoretical entanglements. It usually involves two sets of conditions only, one set that is pertinent to one theory and another set pertinent to an antagonistic theory. It is designed and conducted by a vigorous protagonist of one of the two theories. Now for an experiment to be crucial its design at least must be so accepted by the proponents of each. In practice this rarely occurs. The protagonist of the other theory has his own crucial experiment. In any case the interpreta-

tion of results of a crucial experiment allows plenty of leeway for argument. So crucial experiment follows crucial experiment, sometimes as a sort of parallel play. This represents a distorted system of values in the field of learning. Too much weight is being given to theory. Work on the construction of learning theories is important, and crucial experiments may contribute to theory development, but neither should be valued too highly. To approach our goal of understanding learning, both theory and experimentation are necessary, but the formulation of general laws of learning will depend more upon the results of experiments than upon theories. An experiment in which the independent variable has been varied quantitatively over its entire scale, in which the parameters are identified and measured quantitatively, can result in a quantitative relation between the independent and dependent variables that is susceptible to mathematical expression. This kind of result is a law of learning. It may be a minor law, but it is the stuff from which more general laws of learning will be derived. It is with theory that we must bridge the gap from the minor laws derived directly from experiments to more general laws of learning. In my opinion theories are of secondary importance. If we will do more of the kind of experimentation just mentioned, theory development will pretty much take care of itself.

We will probably always have gaps in our knowledge of learning, and therefore many unsolved problems. It is possible to accept this state of affairs without the necessity of speculation through theory about what is not known. The answer to any given problem will come when someone is ingenious enough to take the problem into the laboratory and solve it by means of experimentation. A theory can be a satisfier, but it is no substitute for knowing by means of experiment.

[MS. received September 18, 1950]

SOME METHODOLOGICAL COMMENTS CONCERNING EXPECTANCY THEORY[1]

BY PAUL E. MEEHL AND KENNETH MacCORQUODALE

University of Minnesota

In attempting to formulate axioms for Tolman's "expectancy" theory we came to certain methodological conclusions which we shall present briefly in the present paper. Space does not permit us to present the twelve postulates thus far developed, nor can we review here the experimental facts. It is our opinion, however, that the recent published and unpublished research is sufficiently favorable toward an expectancy theory to warrant systematic work on its more rigorous formulation.

We early concluded that certain views which have been linked historically to Tolman's formulation are logically unrelated to the "core" concepts of an expectancy theory. Other views seemed related to Tolman's formulation until preliminary efforts to formalize them indicated their independent status as well. We shall merely list those dogmatically as "properties *not* definitive of an expectancy theory": [2]

1. Gestalt-configural stress.
2. Perceptual field stress.
3. Pure contiguity as a sufficient condition.
4. Specification of reaction-class by reference to position, direction, or locomotion (rather than by effector properties).
5. Discontinuity view of discrimination learning.
6. Insistence upon the learning-performance distinction.

What then *is* the "core" concept definitive of an expectancy theory?

According to both S–R theory and to Tolman, the empirical probability of R in the presence of S can be increased in several ways. That more than one way exists is what generates the learning-performance distinction. Thus, without having interpolated further exposures to S, we can raise the strength of R in its presence by starving the rat. But the important point is that operations of this sort are conceived of as acting to change the relation between a theoretically introduced bond sHr and empirical strength. If we examine the manner in which this theoretical bond sHr is *itself* supposed to be strengthened, the three major S–R systems allow for only one kind of operation: ***the running off of the sequence S-then-R*** (or some sequence having sufficient similarity to it to yield generalization and induction effects). The question regarding the *sufficient* conditions for strengthening this sequence, *e.g.*, drive-reduction, is not relevant here. The mere occurrence of the sequence S-then-R may not be sufficient, but all three major S–R systems (Guthrie, Hull, Skinner) make it necessary.

If a postulate is included which allows the theoretical sHr bond to be strengthened by some operation *not* involving the occurrence of the sequence

[1] This is a revision, with very minor changes, of a paper read as part of the Report from the Conference on Behavior Theory (Dartmouth Conference) at the 1950 meeting of the American Psychological Association at State College, Pennsylvania. The other members of the Symposium and Conference are in no way to be considered as endorsing the position presented.

[2] Detailed justification for rejecting these as "essential" features of an expectancy theory will appear in a joint volume to be published by the Dartmouth Conference.

S-then-R (or topographically similar ones), such a single postulate should have striking and diffuse consequences. If we draw up a table of the various experimental designs possible in the single-unit T-maze, and examine those designs in which "positive" outcomes are *prima facie* embarrassing to non-Tolmanites (*i.e.*, whether they hold in common stimulus-reinforcement, drive-reduction, or simple contiguity views), we find that such designs have a common property. This property, which is not found in the designs easily assimilated by S–R theory, is that *during the acquisition phase no increase in the* (subsequently-to-be-manifested) *preference can be observed.* Whenever such an increase is observed during the so-called "latent" phase, the experiment is not disturbing to an S–R theorist. Now if we ask why Tolman is happy with positive outcomes in such designs, we find that he can assume an increase in the sHr bond by an operation *other* than the running off of the S-then-R sequence. Because this operation is sor times taken to be essentially the juxtaposition of two S's, there *is* a certain sense in which Tolman has an "S–S" theory (although the two S's are not necessarily the ones Tolman usually mentions).

Let us suppose, to take a concrete example, that a food-satiated rat runs a T-maze with symmetrical social incentives. No differential strength is observed. Then the hungry rat is *placed* (not having run) in one end-box and allowed to eat. "Appropriate" choice on a subsequent run is *prima facie* embarrassing to non-expectancy theories. It appears that the feeding experience was the operation that strengthened one S.R bond over the other; yet this operation does not involve the occurrence of the sequence S-then-R. This seems to us the real reason that it is embarrassing to non-expectancy theories.

That this situation is what is at stake is further indicated by the way a non-Tolmanite is likely to avoid the embarrassment. He invokes a fractional goal response. That is, he speaks of an accumulation of strength in some bond S.R′ which *does* go on during the latent phase; but the response term R′ is either internal or minimal and hence not noticed by ordinary methods.

Tolman here says what the layman would say: "The feeding experience taught the rat that food was to be found in a certain box, and the latent phase had taught him how to get to that box. So now he tends to go there." Here the strength of sHr is said to increase via the occurrence of two S's, without the occurrence of the sequence S-then-R. Let us call the neutral goal-box stimulation S_2, the "usual terminator" of the sequence S-then-R. The Tolman postulate would seem to involve the assumption that the strength of the S.R bond can be increased *by bringing its usual terminator* S_2 into conjunction with a reinforcer (food, say S*). But this reference to the "usual terminator" can only be understood by thinking of something which the "usual termin*ation*" has set up in the rat. So it seems appropriate to include the usual terminator in one's notation, writing it as "part of" the construct being strengthened during the latent phase. Hence we think in terms of a 3-term designation of the construct, including reference to the eliciting stimulus, the response, and the terminating stimulus. This peculiar "habit" ($S_1R_1S_2$) may be called an "expectancy"; postulates regarding its acquisition, extinction, generalization and "activation" of reaction-potentials are then in order.

Having defined "an expectancy ($S_1R_1S_2$)" implicitly by several postulates which we cannot present here,[3]

[3] The provisional set of postulates is available from the authors upon request.

an expectancy theory seems to involve what might be called a *Postulate of Inference*, as follows:

> The occurrence of a temporal contiguity S_2S^* when the expectancy $(S_1R_1S_2)$ has non-zero strength, produces an increment in the strength of a new expectancy $(S_1R_1S^*)$. The induced strength increases as a decelerated function of the number of such contiguities. The asymptote is the strength of the inducing expectancy $(S_1R_1S_2)$ and the growth rate is a decelerated increasing function of the absolute valence of S^*.

Let us suppose that the experimental facts seem to require such a principle. Does its presence define an expectancy theory? No, because a non-expectancy theorist may be able to derive it as a theorem if suitable changes are made in the formulation of the other postulates as well. The use of r_g by Hullians is, of course, the major case of such a claim. Two comments are in order concerning r_g as a mediator of "inference." First, the derivations tend to be made for special setups, rather than as a general "theorem" to be filled in. Secondly, r_g at present is so poorly quantified a construct that it is as readily available a *deus ex machina* for non-expectancy theorists as the concepts of "attention," "emphasis," or "perceptual threshold" are for Tolman. It is not possible at present to foresee whether the theorem can be derived and can have the combination of quantitative properties in r_g needed to mediate all of the deductions.

Suppose an inference-postulate is derivable as a theorem from "non-expectancy" principles. What non-verbal difference, if any, remains between Tolman and Hull? We believe this question hinges upon the way the word "response" is used. If this word means an *effector*-event, as it originally did to behaviorists, a Tolmanite creates a non-verbal difference by making r_g "central." For this reason we did not list the centralist-peripheralist issue among the pseudo- or non-defining properties of an expectancy theory. Suppose Tolman asserts that the locus of r_g or, better, of the expectancy which plays the same formal rôle as r_g is "in the brain." There are three main reactions to such a claim:

1. It is not part of theory.
2. It is part of theory but I am not interested in it.
3. It is an interesting part of theory but it is false (peripheralism).

There can be no doubt as to the empirical meaningfulness of the centralist thesis. "Where in physical space-time, in what part or phase of the activities occurring within the rat, do the physiological events occur which play the formal rôle in the causal system specified by the sentences of the theory?" One may not care to discuss this thesis, but it is certainly not a transcendental or "meaningless" question. There are respectable sciences of anatomy and physiology which have these events as their subject-matter; and no one but a dualist doubts that these events are the immediate determiners of the movements of an organism. This being the case, it is hard to understand why such a locus-claim is "not *really* part of the theory." A man's theory about any region of inquiry consists of all the empirically meaningful statements that he makes regarding that region. To say that "r_g develops as a growth-function" is theory, whereas to say that "r_g occurs in the effectors" is not theory, seems to us wholly arbitrary. Nor can such a distinction be defended by the methodological dictum that "If two theories have the same consequences they are indistinguishable." The joker here is the phrase "same consequences." That the consequences must be somehow *empirical* is merely the mark of science; but that they must be *behavioral* is a

currently fashionable decision. It may be a good one; but it is only a decision nonetheless!

This brings us to the second attitude, "It's theory but I am not interested in it." Since this is a statement of values, one cannot quarrel with it. But such a value-judgment cannot, by the same token, be imposed on others, *e.g.*, by informing them that the centralist thesis is not part of theory. It has been said that the only reason for constructs is to enable us to "predict behavior." We frankly find it difficult to conceive that people are mainly interested in doing just this. Does anyone but a naturalist have an intrinsic interest in the movements of a rat? We are not erecting theory in order that we may guess whether a rat will turn left or right; on the contrary, we watch which way he turns with bated breath because we wish to confirm our guesses as to the theory. Now this is autobiographical, and others may not share our intrinsic curiosity about the inner events. But those who do not share that curiosity cannot prohibit it in those who do. If a centralist insists that he studies the rat's movements as a *means* to an *end* —to get leads on the character and locus of the inner events—make no mistake about it; he is expressing a very respectable kind of scientific interest. It remains to be seen how far a "purely behavioral" research program can carry him in satisfying such an interest. Let us stress very strongly that *we* have no bets here, and do not intend to desert the T-maze and the Skinner-box for the scalpel. But we do wish to emphasize that any tendency to relegate a "centralism" thesis to the metaphysical limbo upon general *methodological* grounds is doctrinaire. "Purely molar behaviorism" has proved amazingly powerful when expressed as a research preference; but it has no status whatever if treated as if it were a thesis of general epistemology.

Whether the time is ripe for systematic efforts at coördination of molar-behavior constructs with localized inner-events is always problematic. All we can do is to examine critically the evidence for proposed identifications. Whether or not its major contentions are correct, a book such as Hebb's *Organization of Behavior* cannot be dismissed as "irrelevant." It may be tentative, but its neurologizing is not of the bootless and irresponsible sort which we all condemn. Whether it is "premature" remains to be seen. It is our belief that even a purely "molar" research attack is capable of confirming some locus-statements, but not to as high a degree as one could wish. It is, of course, quite possible for a determinedly "molar" behaviorist to ask intelligible questions which he cannot answer (*i.e.*, resolve with any confidence by molar methods). But how far he can go is for future research to decide, and it would be foolish for us to venture any prophecies.

[MS. received for early publication March 8, 1951]

PSYCHOLOGICAL MONOGRAPH

#270

1945

ON PROBLEM SOLVING

By

KARL DUNCKER

$2.50

This popular monograph is #270 of the Psychological Monograph series. It has been reprinted so that it is again available.

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue N. W.

Washington 5, D. C.

American Psychological Association

1515 Massachusetts Ave. N.W.
Washington 5, D. C.

Publications:

AMERICAN PSYCHOLOGIST

Editor: Fillmore H. Sanford, *American Psychological Association.* Contains all official papers of the Association and articles concerning psychology as a profession; monthly.

Subscription: $7.00 (Foreign $7.50). Single copies, $.75.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

Editor: J. McV. Hunt, *Institute of Welfare Research, New York City.* Contains original contributions in the field of abnormal and social psychology, reviews, and case reports; quarterly.

Subscription: $6.00 (Foreign $6.50). Single copies, $1.75.

JOURNAL OF APPLIED PSYCHOLOGY

Editor: Donald G. Paterson, *University of Minnesota.* Contains material covering applications of psychology to business, industry, and education; bi-monthly.

Subscription: $6.00 (Foreign $6.50). Single copies, $1.25.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY

Editor: Harry F. Harlow, *Department of the Army, Washington, D. C.* Contains original contributions in the field of comparative and physiological psychology; bi-monthly.

Subscription: $7.00 (Foreign $7.50). Single copies, $1.25.

JOURNAL OF CONSULTING PSYCHOLOGY

Editor: Laurance F. Shaffer, *Teachers College, Columbia University.* Contains articles in the field of clinical and consulting psychology, counseling and guidance; bi-monthly.

Subscription: $5.00 (Foreign $5.50). Single copies, $1.00.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Editor: Arthur W. Melton, *U. S. Air Force, San Antonio, Texas.* Contains original contributions of an experimental character; monthly, two volumes per year.

Subscription: $14.00 (Foreign $14.50). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

Editor: C. M. Louttit, *University of Illinois.* Contains noncritical abstracts of the world's literature in psychology and related subjects; monthly.

Subscription: $7.00 (Foreign $7.50). Single copies, $.75.

PSYCHOLOGICAL BULLETIN

Editor: Lyle H. Lanier, *University of Illinois.* Contains critical reviews of psychological literature, methodological articles, book reviews, and discussions of controversial issues; bi-monthly.

Subscription: $7.00 (Foreign $7.50). Single copies, $1.25.

PSYCHOLOGICAL MONOGRAPHS: GENERAL AND APPLIED

Editor: Herbert S. Conrad, *U. S. Office of Education, Washington, D. C.* Contains longer researches and laboratory studies which appear as units; published at irregular intervals, about ten numbers per year.

Subscription: $6.00 per volume (Foreign $6.50). Single copies, price varies according to size.

PSYCHOLOGICAL REVIEW

Editor: Carroll C. Pratt, *Princeton University.* Contains original contributions of a theoretical nature; bi-monthly.

Subscription: $5.50 (Foreign $6.00). Single copies, $1.00.